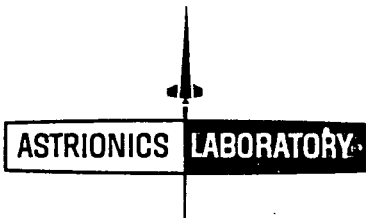
NASA TM X-53044

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

WASHINGTON, D. C.



ASTRIONICS RESEARCH AND DEVELOPMENT REPORT NO. 2

FOR THE PERIOD JULY 1, 1963 THROUGH DECEMBER 31, 1963

ASTRIONICS LABORATORY
RESEARCH AND DEVELOPMENT OPERATIONS
GEORGE C. MARSHALL SPACE FLIGHT CENTER
HUNTSVILLE, ALABAMA

MAY 1, 1964

preface...

This document is the second issue of a Semi-Annual Research and Development Report prepared by the Astrionics Laboratory, Research and Development Operations, George C. Marshall Space Flight Center, Huntsville, Alabama.

The subjects discussed cover technical information that, we believe, will be interesting and helpful to other organizations engaged in space flight and related efforts.

Reprints of these articles may be obtained from the NASA Library, MSFC, Huntsville, Alabama.

W. Haeussermann

W. HAEUSSERMANN
Director, Astrionics Laboratory

ASTRIONICS LABORATORY / 1964

GEORGE C. MARSHALL SPACE FLIGHT CENTER
NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
HUNTSVILLE, ALABAMA

CONTENTS...

I. GUIDANCE AND CONTROL SYSTEMS AND COMPONENTS

Page

A. A Monte Carlo Method of Generating Random State Vectors from the General Multivariate Normal Distribution with an Application to Midcourse Correction by Robert M. Coulter, Robert J. Hill, and Norman C. Perry 1

B. Lunar Descent Guidance Studies by Charles R. MacCrone 12

C. Terminal Guidance Techniques for Soft Lunar Landing by George W. Neal and John F. Pavlick 18

D. Space Vehicle Attitude Control Using Pseudo Rate Modulation by David N. Schultz 24

E. Pilot Participation in Saturn V Guidance and Control by Robert W. Gunderson and Gordon H. Hardy 36

F. Adaptive Phase-Stabilization of Bending for Large Aerospace Vehicle Boosters by Michael T. Borelli and Hans H. Hosenthien 43

G. A System for Acquisition and Reduction of Dynamic Test Data of Large Aerospace Vehicles by Frank Hay 51

H. Signal Conditioning Circuit for Displaying Pitch, Roll, and Yaw on an Oscilloscope by Charles P. Elms 60

I. An Angular Position Converter by Leslie C. Holt 65

II. ELECTRICAL SYSTEM

A. Geometric Analysis of the Mean Normalized Power Available from the Solar Cells on the Micrometeoroid Measurement Capsules by Richard Boehme and Jimmy L. Miller 69

B. Fuel Cell Power for Saturn Instrument Unit by Richard Boehme and Lance Conway 75

III. RESEARCH AND SUPPORTING ENGINEERING

A. Static Load and Flow Equations for a Double-Acting, Journal-Fed, Externally-Pressurized Thrust Bearing by Peter H. Broussard, Jr., John Burch, and Emmett Martz 81

B. An Investigation of Photoelectrically Heterodyning Coherent Light by Cullen Lee 92

C. Thermal Contact Conductance in a Vacuum by Richard M. Dailey and Walter E. Kaspareck 108

D. Flat Conductor Cables by Wilhelm Angele 117

IV. PUBLICATIONS, PRESENTATIONS, AND PATENTS

A. Publications 123
B. Presentations 124
C. Patents 125

acknowledgements...

The articles for this report were contributed by various engineers and physicists of the Astrionics Laboratory, reviewed and compiled by Hubert E. Kroh, and edited by Jeanene Flowers.

Grateful acknowledgement is given to the Technical Publications Section and the Astrionics Support Unit, Graphic Engineering and Model Studies Branch, Management Services Office, MSFC, for preparing the report.

I GUIDANCE AND CONTROL SYSTEMS AND COMPONENTS...

A. A MONTE CARLO METHOD OF GENERATING RANDOM STATE VECTORS FROM THE GENERAL MULTIVARIATE NORMAL DISTRIBUTION WITH AN APPLICATION TO MIDCOURSE CORRECTION

by Robert M. Coulter, Robert J. Hill, and Norman C. Perry *

This article presents a solution to the problem of obtaining random state vectors from the general multivariate normal distribution emanating from several error sources. The method of solution makes use of the theory of multiple correlation.

The method may be applied to many problems involving random error sources. The numerical example chosen for explanation is one involving the midcourse maneuver. A set of random injection errors, which require the midcourse correction, is created from the covariance matrix of injection errors.

INTRODUCTION

The study of the accuracy of guidance hardware can always be performed in terms of probability. In the study of advanced missions such as lunar and interplanetary flights, statistics becomes a particularly useful tool in making comparative analyses of various guidance systems and in determining the figure of merit for a given system.

Linear perturbation theory as applied to statistics has been used in many applications and is discussed in the literature (Ref. 1). A particular application of linear transformations is made (Ref. 2) in assessing the accuracy of the inertial guidance hardware for ascent from earth to injection into a lunar trajectory. The idea here is that if the injection errors are linearly related to the guidance hardware errors, the variances of the respective injection errors are linearly related to the variances of the guidance hardware errors. Thus, a covariance matrix of injection errors can be formed to give a complete knowledge of the errors. This idea is used further (Ref. 3) in computing the midcourse correction velocity requirement caused by the injection errors.

However, when linear theory can no longer be applied, the need arises for other means of handling the statistical analysis of nonlinear functions. This situation occurs when it is desired to know the accuracy with which a vehicle can reach periselenum on a mission such as that of the Lunar Logistic Vehicle. The reason this particular problem is nonlinear is discussed. A midcourse correction is made to null the injection errors from the inertial hardware, but there are errors in the execution of the midcourse correction. Thus, the resultant error, and consequently the periselenum error, is a function of the injection guidance errors as a result of the scale-factor error and the attitude alignment error. The resultant error is also a function of the accelerometer bias and the engine shutdown error, which are independent of the injection errors. The accuracy with which the vehicle reaches periselenum is then a nonlinear function of the injection errors and the execution errors.

All of the errors at midcourse possess statistical distribution. The injection errors are correlated and the execution errors are uncorrelated. A large number of samples can be taken from the respective distributions and the miss at periselenum can be calculated. This is referred to as a Monte Carlo simulation and the results can be expressed in terms of cumulative distributions of miss samples.

*Professor of Mathematics, Auburn University

The sampling from an uncorrelated distribution poses no major problem, but the generation of random numbers from the correlated distribution is not as straightforward. A method is given for selecting random correlated error vector components from a multivariant normal distribution. Tests are described for verifying that the generated vectors are random and that the initial correlations and variances are reproduced in the simulation. An example is given of the application of the simulation; namely, a lunar mission in which the miss at periselenum is considered.

Many other problems may be treated by the Monte Carlo simulation; in fact, this program may be applied to any problem in which the position and velocity errors are given in terms of a covariance matrix from which random errors are to be generated. An example of this would be a rendezvous simulation with the initial errors expressed in a covariance matrix.

METHOD OF GENERATING RANDOM STATE VECTORS

In matrix notation, the multivariate normal distribution $f(\vec{x})$ has the functional representation

$$f(\vec{x}) = \frac{1}{2\pi |\lambda|^{\frac{1}{2}}} e^{-\frac{1}{2}\left|\vec{x}\,\lambda^{-1}\vec{x}^{T}\right|} \tag{1}$$

where the vector $\vec{x}$ is the statistical variable, which is a vector, and λ is the covariance matrix of $\vec{x}$. The 6×6 covariance matrix is shown in equation 2.

$$\lambda = \begin{bmatrix} \sigma_x^2 & \sigma_{xy} & \sigma_{xz} & \sigma_{x\dot{x}} & \sigma_{x\dot{y}} & \sigma_{x\dot{z}} \\ \sigma_{yx} & \sigma_y^2 & \sigma_{yz} & \sigma_{y\dot{x}} & \sigma_{y\dot{y}} & \sigma_{y\dot{z}} \\ \sigma_{zx} & \sigma_{zy} & \sigma_z^2 & \sigma_{z\dot{x}} & \sigma_{z\dot{y}} & \sigma_{z\dot{z}} \\ \sigma_{\dot{x}x} & \sigma_{\dot{x}y} & \sigma_{\dot{x}z} & \sigma_{\dot{x}}^2 & \sigma_{\dot{x}\dot{y}} & \sigma_{\dot{x}\dot{z}} \\ \sigma_{\dot{y}x} & \sigma_{\dot{y}y} & \sigma_{\dot{y}z} & \sigma_{\dot{y}\dot{z}} & \sigma_{\dot{y}}^2 & \sigma_{\dot{y}\dot{z}} \\ \sigma_{\dot{z}x} & \sigma_{\dot{z}y} & \sigma_{\dot{z}z} & \sigma_{\dot{z}\dot{x}} & \sigma_{\dot{z}\dot{y}} & \sigma_{\dot{z}}^2 \end{bmatrix} \tag{2}$$

Let σ_{pq} be any off-diagonal elements, then

$$\sigma_{pq} = \sigma_p \sigma_q \rho_{pq} = \sigma_{qp}. \tag{3}$$

Notice that the covariance matrix describes all of the statistical parameters of the multivariate normal distribution. The diagonal terms are the variances of the components of $\vec{x}$. The value ρ_{pq} in equation 3 is the correlation between the components p and q. ρ_{pq} may be obtained by equation 4.

$$\rho_{pq} = \frac{\sigma_{pq}}{\sigma_p \sigma_q} = \rho_{qp}. \tag{4}$$

Also note that $\rho_{pp} = 1$. The matrix of the correlation coefficients is called the correlation matrix and is denoted by R.

$$R = \begin{bmatrix} 1 & \rho_{xy} & \rho_{xz} & \rho_{x\dot{x}} & \rho_{x\dot{y}} & \rho_{x\dot{z}} \\ \rho_{yx} & 1 & \rho_{yz} & \rho_{y\dot{x}} & \rho_{y\dot{y}} & \rho_{y\dot{z}} \\ \rho_{zx} & \rho_{zy} & 1 & \rho_{z\dot{x}} & \rho_{z\dot{y}} & \rho_{z\dot{z}} \\ \rho_{\dot{x}x} & \rho_{\dot{x}y} & \rho_{\dot{x}z} & 1 & \rho_{\dot{x}\dot{y}} & \rho_{\dot{x}\dot{z}} \\ \rho_{\dot{y}x} & \rho_{\dot{y}y} & \rho_{\dot{y}z} & \rho_{\dot{y}\dot{x}} & 1 & \rho_{\dot{y}\dot{z}} \\ \rho_{\dot{z}x} & \rho_{\dot{z}y} & \rho_{\dot{z}z} & \rho_{\dot{z}\dot{x}} & \rho_{\dot{z}\dot{y}} & 1 \end{bmatrix} \tag{5}$$

The correlation matrix is necessary for the derivation of later equations.

To obtain a random vector, $(x, y, z, \dot{x}, \dot{y}, \dot{z})$, the component x is first selected from a normal population of zero mean and standard deviation σ_x.

For convenience, assume that a random number may be generated from a normal population of zero mean and unit variance. Statistical tests are described and performed later to justify this assumption and demonstrate the reliability of such random number selectors or generators.

The first step in generating the six correlated components of the random vector is to select one of the components to be random according to its mean and standard deviation. In this discussion, x will be chosen to be this free variable and the other five components will be chosen with the constraints of the intercorrelations. The results would be the same if any one of the components of the random vector were chosen as the free variable.

Let t be a random number selected from the normal population with $\mu = 0$ and $\sigma = 1$; then the transformation

$$t = \frac{x - \mu_x}{\sigma_x} \tag{6}$$

may be used to obtain x.

Now that x has been chosen, there is not the same freedom to select y since x and y are correlated. The components x and y are related by the coefficient of correlation ρ_{xy}. The equation of the regression line of y on x is

$$y = x\, \rho_{xy} \frac{\sigma_y}{\sigma_x}. \tag{7}$$

The equation of the regression line may be obtained by applying the method of least squares.

The expected value of y, given a value of x, is readily obtained by substituting x into equation 7. This expected value of y is the mean of y, given a value of x, and is denoted by the symbol μ_y^*. The standard deviation of y, given a value of x, will be denoted by σ_y^* and is given in equation 8,

$$\sigma_y^* \quad \sigma_y\, (1-\rho_{xy}^2)^{\frac{1}{2}}, \tag{8}$$

which represents the variability of y about the regression line of y on x. Note that the expression for σ_y^* is independent of the value selected for x. A measure of the concentration of points about the regression line is given by the expression $\frac{1}{N} \sum_{i=n}^{n} D_i^2$ where D_i is the difference between the expected value of y and the observed value of y. The value $\frac{1}{N} \sum_{i=1}^{n} D_i^2$ is called the square of the standard error of estimate and is denoted by S_y^2.

$$S_y^2 = \frac{1}{N} \Sigma\, (y - \rho_{xy} \frac{\sigma_y}{\sigma_x} x)^2 \tag{9}$$

$$NS_y^2 = \Sigma\, y^2 - 2\, \rho_{xy} \frac{\sigma_y}{\sigma_x} \Sigma xy + \rho_{xy}^2 \frac{\sigma_y^2}{\sigma_x^2} \Sigma\, x^2 \tag{10}$$

$$= N\, \sigma_y^2 - 2\, n\, \rho_{xy}^2\, \sigma_y^2 + N\, \rho_{xy}^2\, \sigma_y^2 \tag{11}$$

$$= N\, \sigma_y^2\, (1 - \sigma_{xy}^2) \tag{12}$$

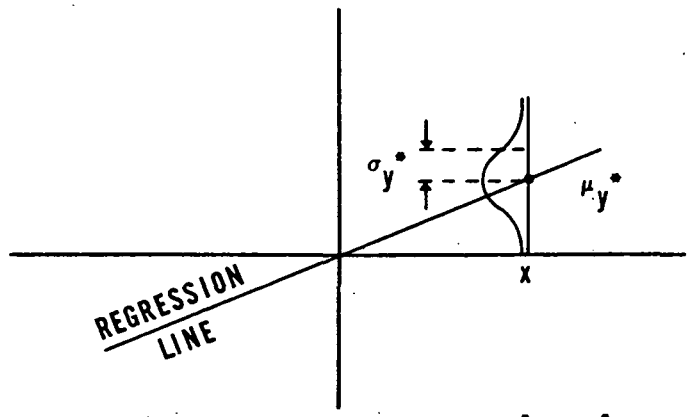


FIGURE 1. GEOMETRIC REPRESENTATION OF μ_y^* AND σ_y^*

$$S_y = \sigma_y\, (1 - \sigma_{xy}^2)^{\frac{1}{2}}. \tag{13}$$

The two results μ_y^* and σ_y^* are displayed geometrically in Figure 1.

To obtain y, a random number t is selected from the normal population with zero mean and unit variance and the transformation

$$y = t\, \sigma_y^* + \mu_y^* \tag{14}$$

is used.

Now that x and y have been chosen, the freedom to select z is even more restricted since z and x are correlated. The value selected for z will be a function of x, y, σ_x, σ_y, σ_z, ρ_{xy}, ρ_{xz}, and ρ_{yz}, all of which are known.

The following notation is introduced to explain the following equations. Define R_k as the k × k determinant or subdeterminant of the correlation matrix, R, which contains the k^{th} row and column and every row and column designated with subscripts less than k. For example,

$$R_2 = \begin{vmatrix} \rho_{11} & \rho_{12} \\ \rho_{21} & \rho_{22} \end{vmatrix} = \begin{vmatrix} \rho_{xx} & \rho_{xy} \\ \rho_{xy} & \rho_{yy} \end{vmatrix} = \begin{vmatrix} 1 & \rho_{xy} \\ \rho_{xy} & 1 \end{vmatrix} \tag{15}$$

$$R_3 = \begin{vmatrix} \rho_{11} & \rho_{12} & \rho_{13} \\ \rho_{21} & \rho_{22} & \rho_{23} \\ \rho_{31} & \rho_{32} & \rho_{33} \end{vmatrix} \tag{16}$$

The determinants are symmetrical and the diagonal elements are equal to one. Define R_{ij} as the cofactor of ρ_{ij} in the determinant R_j.

For example,

$$R_{14} = \begin{vmatrix} \rho_{21} & \rho_{22} & \rho_{23} \\ \rho_{31} & \rho_{32} & \rho_{33} \\ \rho_{41} & \rho_{42} & \rho_{43} \end{vmatrix} \tag{17}$$

Note that the cofactor of ρ_{ij} carries the sign $(-1)^{i+j}$.

With the preceding notation, the equation of the regression plane of z on x and y is given by

$$\frac{x}{\sigma_x} R_{zx} + \frac{y}{\sigma_y} R_{zy} + \frac{z}{\sigma_z} R_{zz} = 0. \tag{18}$$

It follows from equation 18 that the expected value of z, which is μ_z^*, is obtained by substituting the values obtained for x and y into equation 19.

$$\mu_z^* = \frac{-\sigma_z}{R_{zz}} \left(\frac{x}{\sigma_x} R_{zx} + \frac{y}{\sigma_y} R_{zy}\right). \tag{19}$$

The standard deviation of z, given x and y, is derived by the expression

$$\sigma_z^* = \sigma_z \left(\frac{R_3}{R_{zz}}\right)^{\frac{1}{2}}. \tag{20}$$

Notice that σ_x^* is independent of the values of x and y. To derive the expression for the variance of z, given two other variables, let D be the difference between the observed and expected value of z.

$$N S_z^2 = \sum_{i=1}^{n} D_i^2 \tag{22}$$

$$N S_z^2 = \Sigma \left\{z - \left[-\frac{\sigma_z}{R_{zz}} \left(\frac{x}{\sigma_x} R_{zx} + \frac{y}{\sigma_y} R_{zy}\right)\right]\right\}^2. \tag{22}$$

Equation 22, as in Reference 4, may be reduced to

$$\sigma_z^* = \sigma_z \left(\frac{R_3}{R_{zz}}\right)^{\frac{1}{2}}. \tag{23}$$

The results are represented geometrically in Figure 2.

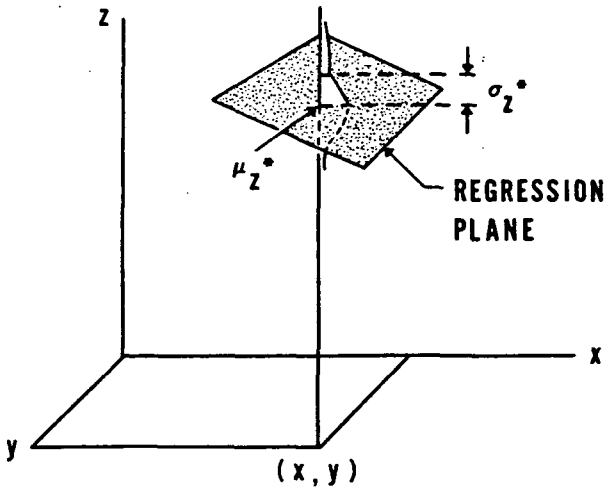


FIGURE 2. GEOMETRIC REPRESENTATION OF μ_z^* AND σ_z^*

Given μ_z^* and σ_z^* and a random number t from the normal population with zero mean and unit variance, the transformation

$$t = \frac{z - \mu_x^*}{\sigma_z^*} \tag{24}$$

may be used to obtain a value of z.

The formulas for obtaining $\dot{x}$, $\dot{y}$, and $\dot{z}$ are derived in a similar manner.

$$\mu_{\dot{x}}^* = \frac{-\sigma_{\dot{x}}}{R_{\dot{x}\dot{x}}} \left(\frac{x}{\sigma_x} R_{\dot{x}x} + \frac{y}{\sigma_y} R_{\dot{x}y} \quad \frac{z}{\sigma_z} R_{\dot{x}z}\right) \tag{25}$$

$$\sigma_{\dot{x}}^* = \sigma_{\dot{x}} \left(\frac{R^4}{R_{\dot{x}\dot{x}}}\right)^{\frac{1}{2}} \tag{26}$$

$$\dot{x} = t\,\sigma_{\dot{x}}^* + \mu_{\dot{x}}^* \tag{27}$$

$$\mu_{\dot{y}}^* = \frac{-\sigma_{\dot{y}}}{R_{\dot{y}\dot{y}}} \left(\frac{x}{\sigma_x} R_{\dot{y}x} + \frac{\dot{y}}{\sigma_y} R_{\dot{y}y} + \frac{z}{\sigma_{\dot{x}}} R_{yz} + \frac{\dot{x}}{\sigma_{\dot{x}}} R_{\dot{y}\dot{x}}\right) \tag{28}$$

$$\sigma_{\dot{y}}^* = \sigma_{\dot{y}} \left(\frac{R_5}{R_{\dot{y}\dot{y}}}\right)^{\frac{1}{2}} \tag{29}$$

$$\dot{y} = t\,\sigma_{\dot{y}}^* + \mu_{\dot{y}}^*. \tag{30}$$

$$\mu_{\dot{z}}^* = \frac{-\sigma_{\dot{z}}}{R_{\dot{z}\dot{z}}} \left(\frac{x}{\sigma_x} R_{\dot{z}x} + \frac{y}{\sigma_y} R_{\dot{z}y} + \frac{z}{\sigma_z} R_{\dot{z}z} + \frac{\dot{x}}{\sigma_{\dot{x}}} R_{\dot{z}\dot{x}} + \frac{\dot{y}}{\sigma_{\dot{y}}} R_{\dot{z}\dot{y}}\right) \quad (31)$$

$$\sigma_{\dot{z}}^* = \sigma_{\dot{z}} \left(\frac{R_6}{R_{\dot{z}\dot{z}}}\right)^{\frac{1}{2}} \quad (32)$$

$$\dot{z} = t\,\sigma_{\dot{z}}^* + \mu_{\dot{z}}^* . \quad (33)$$

The value t in equations 27, 30, and 33 represents random numbers from the normal population with zero mean and unit variance.

With this method, random numbers from multi-dimensional normal distributions may be generated. Here the interest was with a six dimensional vector; however, the number of variables could be increased or decreased by extending or truncating the method.

SELECTION OF RANDOM NUMBERS

The speed of modern computing machines has made possible the utilization of the Monte Carlo methods. In general, these methods make use of a statistical distribution that describes a phenomenon to be studied. Computers can generate large random samples that have the properties of described distributions.

The basic need of the method previously described is random numbers from the normal population with zero mean and unit variance. This discussion is concerned with statistical tests to determine that the numbers are random and that they come from the desired normal population. Without meeting these requirements, the method of obtaining correlated random numbers cannot possibly be successful. It is also demonstrated that the correlated random numbers are generated from the desired population.

To satisfy the need of the driving routine, i.e., generation of random numbers from the normal population of zero mean and unit variance, a random number generator was selected. Several statistical tests were performed and the results were satisfactory.

The χ^2 test is the most common to perform in checking a statistical distribution. The principle of the χ^2 test is to compare expected results with the obtained results. Ten runs were made and the values of χ^2 obtained were: 23.461, 36.713, 30.062, 40.096, 34.871, 30.010, 26.522, 23.263, and 19.655. The critical value of χ^2 leading to the rejection of the hypothesis that the numbers are normally distributed is $\chi_0^2 = 42.557$. The value of χ_0^2 is based on a 5 percent critical value. Figure 3, a plot of the distribution of the χ^2's obtained and the theoretical curve, shows that the values obtained for χ^2 approximate the theoretical curve. Since all of the sample values are less than 42.557, the hypothesis that the population is normal must be accepted.

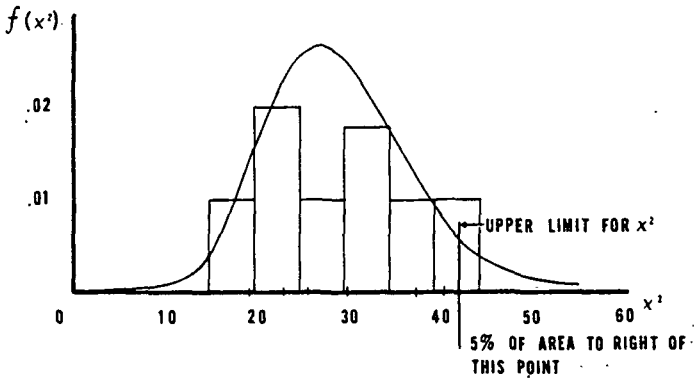


FIGURE 3. THE EMPIRICAL AND THEORETICAL DISTRIBUTION OF THE χ^2's

The next test was performed to check the distribution of the mean. The mean of a random sample possesses a normal distribution with $\mu_{\bar{x}} = \mu$ and $\sigma_{\bar{x}} = \sigma/\sqrt{N}$ where N is the sample size. Ten runs were made, each with a sample size of 10,000. Values of the means were: -.00552, .00540, -.00142, .00388, -.000575, .000150, -.00952, .0142, -.00855, and .0118. The critical value for rejecting the hypothesis that $\mu = 0$ is $\bar{x} = \pm .02$. Figure 4 shows the distribution of the 10 means with the theoretical probability distribution.

All of the sample means were well within the desired bounds (Fig. 4); therefore, the hypothesis that $\mu = 0$ must be accepted.

The next test performed on the driving routine was to check the distribution of the variance which, theoretically, should be one. If the variance S^2 is

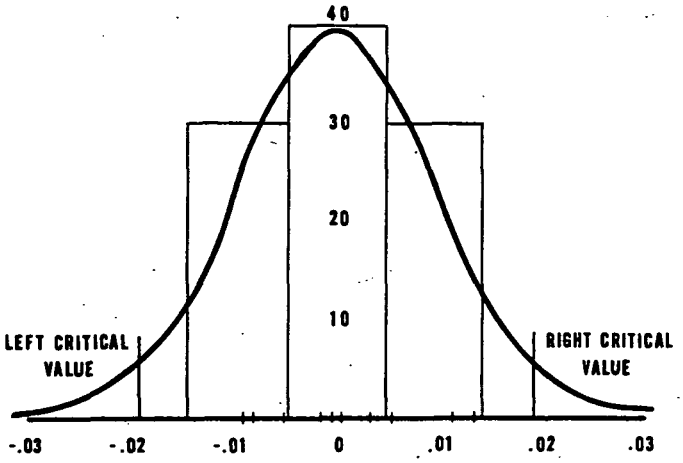


FIGURE 4. DISTRIBUTION OF MEANS

computed from a sample size N from a normal population with variance σ^2, then the statistic NS^2/σ^2 possesses a χ^2 distribution with N - 1 degrees of freedom. Therefore, to test the hypothesis that $\sigma^2 = 1$, the equation

$$\frac{NS^2}{\sigma^2} = \chi^2 \tag{34}$$

is computed and compared to the χ^2 distribution with 9999 degrees of freedom since the sample size used was 10,000. The values of NS/σ^2 that were obtained from 10 samples of 10,000 each are: 10029.2, 9918.3, 10169.0, 10041.4, 9867.9, 10161.6, 9783.6, 9907.1, 9903.8, 9978, and 8625. The critical values of NS^2/σ^2 for rejecting the hypothesis that $\sigma^2 = 1$ are $9578.5 < NS^2/\sigma^2 < 10,427.2$. Notice that all of the values are within the desired range, thus the hypothesis that $\sigma^2 = 1$ must be accepted.

A nonparametric test was also made on the sample to check the randomness. The number of runs possesses a definite statistical distribution, depending on the size of the sample. The variable τ possesses a normal distribution with zero mean and unit variance with

$$\tau = \frac{\nu - \mu_\nu + \frac{1}{2}}{\sigma_\nu} \tag{35}$$

where ν is the number of runs and

$$\mu_\nu = \frac{2 N_a N_b}{N_a + N_b} + 1 \tag{36}$$

and

$$\sigma_\nu^2 = \frac{2 N_a N_b (2 N_a N_b - N_a - N_b)}{(N_a + N_b)^2 (N_1 + N_2 - 1)} \tag{37}$$

where N_a is the number of sample values above the median and N_b is the number of samples below the median.

Five samples of 1,000 numbers each were run from the driving routine. The following number of runs were obtained: 490, 475, 494, 492, and 530. The critical value for rejecting the hypothesis is $469 < \nu < 532$. Since all of these values are within limits, the hypothesis that the samples are random, based on the theory of runs, must be accepted.

The remaining test to check the sample for randomness was also nonparametric. This test applies the theory of serial correlation. Let the values $x_1, x_2, x_3, \ldots, x_n$ be the order in which a sample of size N was generated. The statistic

$$R = \sum_{i=1}^{n-1} x_i x_{i+1} \tag{38}$$

is the serial correlation coefficient. This value is distributed normally with

$$\mu = \frac{S_1^2 - S_2^2}{N - 1}$$

$$\sigma^2 \quad \frac{S_2^2 - S_1^4}{N - 1} + \frac{S_1^4 - 4S_1^2 S_2 + 4S_1 S_3 + S_2^2 - 2S_4}{(N - 1)(N - 2)} - \mu^2 \tag{39}$$

where

$$s_\nu = \sum_{j=1}^{n} x_j^\nu$$

$$\nu = 1, 2, 3, 4. \tag{40}$$

To apply this test, 5 samples of 100 numbers each were used to compute the statistic R along with its mean and variance. The following values were obtained for R:

$R_1 = 54.62$	$\mu_1 = 58.86$	$\sigma_1 = 4.003$
$R_1 = 73.65$	$\mu_2 = 74.79$	$\sigma_2 = 4.006$
$R_3 = 55.48$	$\mu_3 = 54.48$	$\sigma_3 = 3.164$
$R_4 = 78.17$	$\mu_4 = 76.67$	$\sigma_4 = 4.514$
$R_5 = 67.18$	$\mu_5 = 62.72$	$\sigma_5 = 2.905$

All of the R's are within the two-sigma significance bound of the mean. Thus the hypothesis that the samples are random must be accepted.

On the basis of all of the tests, the driving routine must be considered to generate random numbers from the normal population with zero mean and unit variance.

With the existence of such a random number generator, the generation of random vectors from the desired multivariable normal population is possible by using the formulas previously derived. To ascertain how well the random vectors approximate the desired distribution, the parameters of the multivariate distribution may be computed from a large sample of random vectors. For example, a set of 5000 random vectors was generated from the population represented in

$$\lambda = \begin{bmatrix} 79000. & 7411. & 0 & 237.7 & 92.09 & 0 \\ 7411. & 9414. & 0 & 28.47 & 26.40 & 0 \\ 0 & 0 & 94143. & 0 & 0 & 225.9 \\ 237.7 & 28.47 & 0 & 0.7378 & 0.3006 & 0 \\ 92.09 & 26.40 & 0 & 0.3006 & 0.1592 & 0 \\ 0 & 0 & 225.9 & 0 & 0 & 0.60903 \end{bmatrix} \quad (41)$$

The population is described by fifteen distinct parameters that were computed from the sample using the maximum likelihood estimators of each of these values as listed in Table I.

To further study the joint distribution of the six variables, the correlation matrix is computed for each of the 10 samples of random vectors. The statistical distribution of the coefficient of correlation is known. The variable υ is normally distributed if

$$\upsilon = \tfrac{1}{2} \ln \frac{1 + r}{1 - r} \quad (42)$$

with

$$\mu_\upsilon = \tfrac{1}{2} \ln \frac{1 + \rho}{1 - \rho} \quad (43)$$

and

$$\sigma_\upsilon = \frac{1}{N - 3} \quad (44)$$

where r is the coefficient of correlation computed from a sample of size N, and ρ is the population coefficient of correlation. Notice that

$$r = \tanh \upsilon. \quad (45)$$

The value υ may be reduced to normal units by the transformation

$$\tau = \frac{\upsilon - \mu_\upsilon}{\sigma_\upsilon}. \quad (46)$$

If the two-sigma value is set for the bounds of r,

$$\pm 2\,\sigma_\upsilon + \mu_\upsilon = \upsilon. \quad (47)$$

Thus r is acceptable if

$$\tanh (-2\,\sigma_\upsilon + \mu_\upsilon) < r < \tanh (2\sigma_\upsilon + \mu_\upsilon). \quad (48)$$

TABLE I

σ_x^2 =	79311. m^2	$\sigma_{y\dot{z}}$ =	1,468 m^2/s
σ_{xy} =	8419. m^2	σ_z^2 =	94000 m^2
σ_{xz} =	-298. m^2	$\sigma_{z\dot{x}}$ =	0.4664 m^2/s
$\sigma_{x\dot{x}}$ =	239.5 m^2/s	$\sigma_{z\dot{y}}$ =	1.631 m^2/s
$\sigma_{x\dot{y}}$ =	94.68 m^2/s	$\sigma_{z\dot{z}}$ =	225.1 m^2/s
$\sigma_{x\dot{z}}$ =	-1.079 m^2/s	$\sigma_{\dot{x}}^2$ =	0.7458 m^2/s^2
σ_y^2 =	9643 m^2	$\sigma_{\dot{x}\dot{y}}$ =	0.3104 m^2/s^2
σ_{yz} =	509.6 m^2	$\sigma_{\dot{x}\dot{z}}$ =	0.0006 m^2/s^2
$\sigma_{y\dot{x}}$ =	31.88 m^2/s	$\sigma_{\dot{y}}^2$ =	0.1664 m^2/s^2
$\sigma_{y\dot{y}}$ =	28.17 m^2/s	$\sigma_{\dot{y}\dot{z}}$ =	0.0050 m^2/s^2
	$\sigma_{\dot{z}}^2$ = 0.6062 m^2/s^2		

TABLE II

ρ_{xy} =	.2929	$\rho_{y\dot{y}}$ =	.6988
ρ_{xz} =	-.0198	$\rho_{y\dot{z}}$ =	.0054
$\rho_{x\dot{x}}$ =	.9848	$\rho_{z\dot{x}}$ =	-.0144
$\rho_{x\dot{y}}$ =	.8195	$\rho_{z\dot{y}}$ =	-.0079
$\rho_{x\dot{z}}$ =	-.0266	$\rho_{z\dot{z}}$ =	.9447
ρ_{yz} =	.0046	$\rho_{\dot{x}\dot{y}}$ =	.8760
$\rho_{y\dot{x}}$ =	.3632	$\rho_{\dot{x}\dot{z}}$ =	-.0216
	$\rho_{\dot{y}\dot{z}}$ = -.0126		

Table II is a listing of the 15 computed coefficients of correlations. The limits on these values are:

.2436	$<\rho_{xy}<$	.2977
-.0282	$<\rho_{xz}<$	.0282
.9838	$<\rho_{x\dot{x}}<$	.9855
.8116	$<\rho_{x\dot{y}}<$	.8301
-.0282	$<\rho_{x\dot{z}}<$	.0282
-.0282	$<\rho_{yz}<$	.0282
.3163	$<\rho_{y\dot{x}}<$	.3663
.6666	$<\rho_{y\dot{y}}<$	.6969
-.0282	$<\rho_{y\dot{z}}<$	.0282
-.0282	$<\rho_{z\dot{x}}<$	.0282
-.0282	$<\rho_{z\dot{y}}<$	.0282
.9402	$<\rho_{z\dot{z}}<$	.9464
.8702	$<\rho_{\dot{x}\dot{y}}<$	.8833
-0282	$<\rho_{\dot{x}\dot{z}}<$	.0282
-.0282	$<\rho_{\dot{y}\dot{z}}<$	.0282

An examination of Table II shows that only 1 of the 15 values is out of bounds and confirms that the values came from the desired population.

NUMERICAL EXAMPLE

The ability to select a set of random vector components from a covariance matrix can be usefully applied in the study of the accuracy of inertial guidance systems which might be used on future lunar and interplanetary missions.

The inertial guidance system under consideration can be assumed to have "n" independent, normally-distributed error sources (accelerometer scale factor, gyro drift, etc.). By known methods (Refs. 1 and 2), these errors can be transformed into a covariance matrix of injection errors and then into a covariance matrix of midcourse correction components. Thus, the statistics of injection errors and midcourse corrections can be obtained.

The same inertial guidance system, however, will introduce random errors into the execution of the midcourse maneuver. Since a random set of injection error components can be selected, a method of simulating the effect of a given set of injection errors on the position error at the target is proposed. The simulation includes tracing the effect of the injection error through several midcourse corrections, each with random execution errors.

The philosophy of the midcourse correction is to make a first correction to null the position error at the nominal time of target arrival. The second maneuver will also correct the position error at target arrival, this time correcting the execution error of the first correction.

The execution errors used in the study reflect errors in attitude as well as errors in magnitude of the correction; the four considered are bias, scale factor, and two pointing angle errors; all are assumed to be normally distributed and independent.

This discussion is based on the theory of linear perturbation. Reference 3 briefly describes the transformation matrices used in the equations, as well as the methods of generating them.

From the covariance matrix of injection errors, a random set of error components is selected as follows: x_o, y_o, z_o, $\dot{x}_o$, $\dot{y}_o$, $\dot{z}_o$ (the delta notation is

dropped for simplicity). The following equation describes how the error is transformed into the components of velocity to be added at the first correction time ($\dot{x}_{c1}$, $\dot{y}_{c1}$, $\dot{z}_{c1}$):

$$\begin{bmatrix} \dot{x}_{c1} \\ \dot{y}_{c1} \\ \dot{z}_{c1} \end{bmatrix} = \begin{bmatrix} \frac{\partial \dot{x}_{c1}}{\partial x_o} & \frac{\partial \dot{x}_{c1}}{\partial y_o} & \frac{\partial \dot{x}_{c1}}{\partial z_o} & \frac{\partial \dot{x}_{c1}}{\partial \dot{x}_o} & \frac{\partial \dot{x}_{c1}}{\partial \dot{y}_o} & \frac{\partial \dot{x}_{c1}}{\partial \dot{z}_o} \\ \frac{\partial \dot{y}_{c1}}{\partial x_o} & \frac{\partial \dot{y}_{c1}}{\partial y_o} & \frac{\partial \dot{y}_{c1}}{\partial z_o} & \frac{\partial \dot{y}_{c1}}{\partial \dot{x}_o} & \frac{\partial \dot{y}_{c1}}{\partial \dot{y}_o} & \frac{\partial \dot{y}_{c1}}{\partial \dot{z}_o} \\ \frac{\partial \dot{z}_{c1}}{\partial x_o} & \frac{\partial \dot{z}_{c1}}{\partial y_o} & \frac{\partial \dot{z}_{c1}}{\partial z_o} & \frac{\partial \dot{z}_{c1}}{\partial \dot{x}_o} & \frac{\partial \dot{z}_{c1}}{\partial \dot{y}_o} & \frac{\partial \dot{z}_{c1}}{\partial \dot{z}_o} \end{bmatrix} \begin{bmatrix} x_o \\ y_o \\ z_o \\ \dot{x}_o \\ \dot{y}_o \\ \dot{z}_o \end{bmatrix} \quad (49)$$

where the generation of the transformation matrix is described in Reference 2.

The results of the first equation, i.e., the components of the first correction velocity, can be acted upon by a random selection of the execution errors and transformed into a set of error components of the maneuver ($\Delta\dot{x}_{c1}$, $\Delta\dot{y}_{c1}$, $\Delta\dot{z}_{c1}$).

Instead of the first maneuver nulling the position error at the intended target, the error components ($\Delta\dot{x}_{c1}$, $\Delta\dot{y}_{c1}$, $\Delta\dot{z}_{c1}$) will cause a position deviation at the target shown in the following equation as x_t, y_t, z_t:

$$\begin{bmatrix} x_t \\ y_t \\ z_t \end{bmatrix} = \begin{bmatrix} \frac{\partial x_t}{\partial \dot{x}_1} & \frac{\partial x_t}{\partial \dot{y}_1} & \frac{\partial x_t}{\partial \dot{z}_t} \\ \frac{\partial y_t}{\partial \dot{x}_1} & \frac{\partial y_t}{\partial \dot{y}_1} & \frac{\partial y_t}{\partial \dot{z}_1} \\ \frac{\partial z_t}{\partial \dot{x}_1} & \frac{\partial z_t}{\partial \dot{y}_1} & \frac{\partial z_t}{\partial \dot{z}_1} \end{bmatrix} \begin{bmatrix} \Delta\dot{x}_{c1} \\ \Delta\dot{y}_{c1} \\ \Delta\dot{z}_{c1} \end{bmatrix} \quad (50)$$

At some later time, a second correction velocity vector can be computed as follows:

$$\begin{bmatrix} \dot{x}_{c2} \\ \dot{y}_{c2} \\ \dot{z}_{c2} \end{bmatrix} = \begin{bmatrix} \frac{\partial x_t}{\partial \dot{x}_2} & \frac{\partial x_t}{\partial \dot{y}_2} & \frac{\partial x_t}{\partial \dot{z}_2} \\ \frac{\partial y_t}{\partial \dot{x}_2} & \frac{\partial y_t}{\partial \dot{y}_2} & \frac{\partial y_t}{\partial \dot{z}_2} \\ \frac{\partial z_t}{\partial \dot{x}_2} & \frac{\partial z_t}{\partial \dot{y}_2} & \frac{\partial z_t}{\partial \dot{z}_2} \end{bmatrix}^{-1} \begin{bmatrix} x_t \\ y_t \\ z_t \end{bmatrix} \quad (51)$$

Once again, the random execution errors will act upon this maneuver and yield a second set of execution error components ($\Delta\dot{x}_{c2}$, $\Delta\dot{y}_{c2}$, $\Delta\dot{z}_{c2}$). A new position deviation is obtained (x_t', y_t', z_t') by changing the subscripts in equation 50.

To demonstrate the use of the preceding formulation, numerical results that reflect a typical lunar mission are shown.

The covariance matrix of injection errors (equation 41) gives the magnitude of injection error that might result from a typical inertial guidance system acting through a first burn, coast period, and second burn.

The sets of transformation matrices reflect a 72-hour flight time from injection to periselenum and two midcourse corrections, the first occurring at 10 hours and the second at 56 hours.

The matrices were generated with a computer program using the restricted three-body model (Ref. 5). The coordinate system is described in Figure 5.

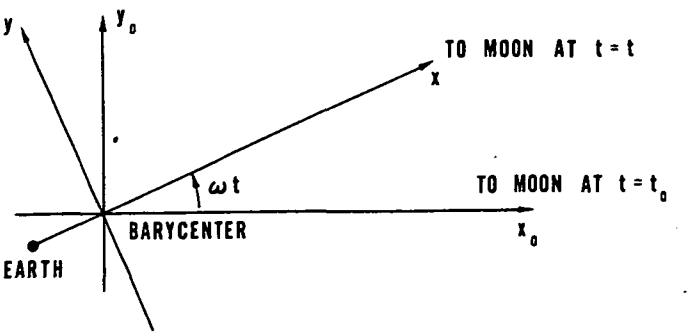


FIGURE 5. COORDINATE SYSTEM OF RESTRICTED THREE-BODY MODEL

All execution errors are assumed to have normal independent distributions with means equaling zero. The scale factor standard deviation is assumed to be 0.033 percent, the bias standard deviation is assumed to be 0.033 m/s, and the pointing angle error standard deviations are both equal to 0.167 degrees.

Using the preceding data, the results of the Monte Carlo simulation are shown in Figures 6 through 11. These figures describe the frequency of occurrence versus magnitude of the quantity being observed. The results are from a sample of 10,000 random injection errors.

Note that the following results depend upon the particular covariance matrix under consideration.

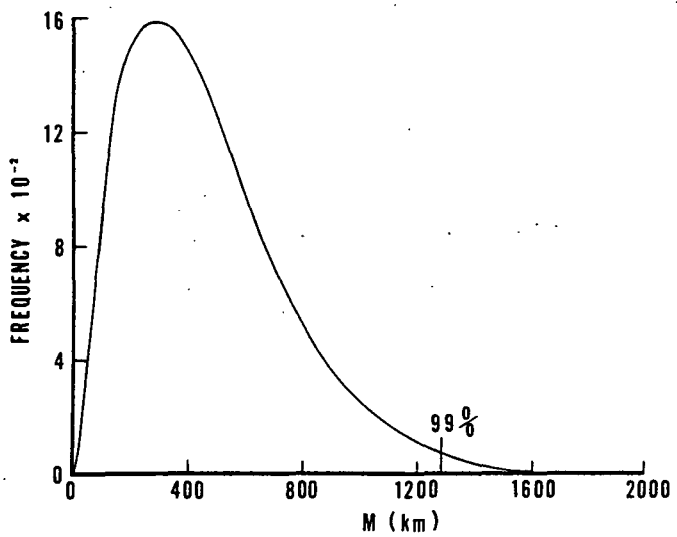


FIGURE 6. FREQUENCY VERSUS UNCORRECTED MISS (M)

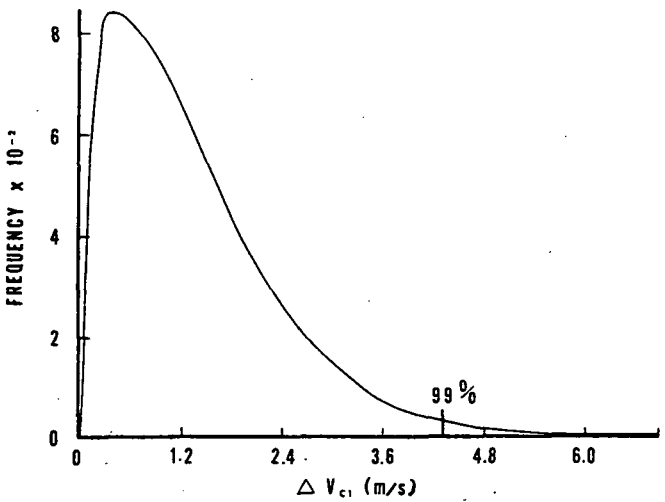


FIGURE 7. FREQUENCY VERSUS ΔV_{C1}

The results shown in Figure 6 give the size of the uncorrected miss (M) at the target. This graph gives the frequency of position miss at the target if no midcourse corrections are initiated.

Figure 7 shows the first midcourse velocity increment (ΔV_{C_1}) to be less than 4.3 m/s in 99 percent of the 10,000 cases considered. The result is compatible with results obtained by the statistical approach described in Reference 4.

Figure 8 describes the distribution of the target miss (M_1) remaining after the first correction if no second correction is made. This miss is the amount of error attributed to the execution errors in the first maneuver.

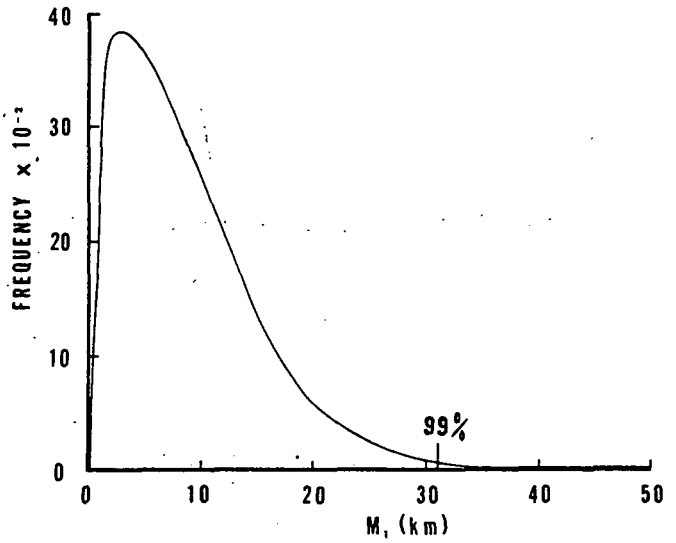


FIGURE 8. FREQUENCY VERSUS M_1

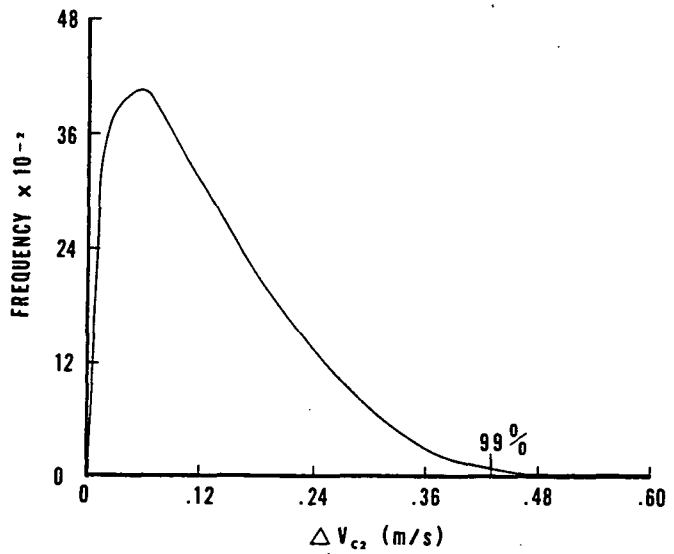


FIGURE 9. FREQUENCY VERSUS ΔV_{C2}

The results of the magnitude of the second correction (ΔV_{C_2}) are given in Figure 9. The indication is that an allowance of $\frac{1}{2}$ m/s is sufficient for the second correction.

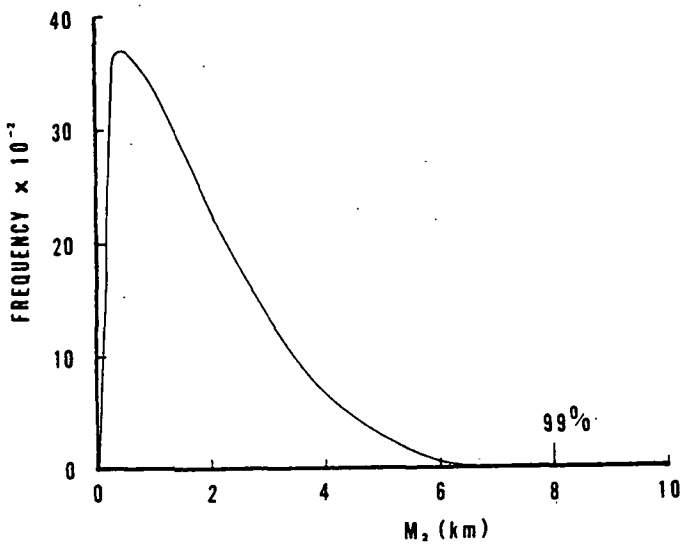


FIGURE 10. FREQUENCY VERSUS M_2

In Figure 10, the frequency of the remaining target error (M_2) is shown again. This time, however, it shows the distribution of the target miss after the second correction, assuming the second correction is not perfectly executed.

Results of further simulation, holding some of the errors equal to zero, indicate virtually all of the target error is caused by bias. Ten thousand cases where scale factor was the only error source showed negligible target error. The pointing angle errors, when considered alone, also give negligible errors. The effect of bias error (M_2'), when considered alone, is shown in Figure 11.

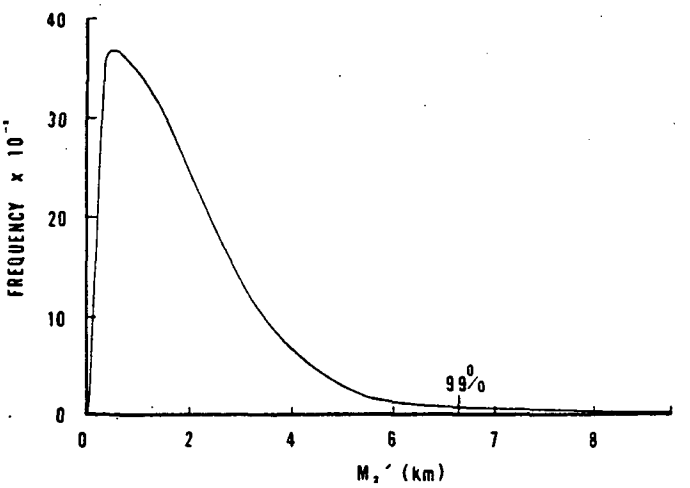


FIGURE 11. FREQUENCY VERSUS M_2'

REFERENCES

1. Noton, A. R. Maxwell, "The Statistical Analysis of Space Guidance Systems," Jet Propulsion Laboratory, C. L. T., TM Nr. 33-15, June 15, 1960.

2. Neighbors, Alice K.,; Cole, John; Daniel, Fred; "Error Analysis of Saturn Guidance Hardware as Applied to a Lunar Mission," MTP-ASTR-A-63-4, March 13, 1963.

3. Coulter, Robert M., "Midcourse Correction Requirements for the Lunar Logistics Vehicle," NASA TM X-53,000, October 1, 1963

4. Kenney, John F., Mathematics of Statistics, Part 2, D. Van Nostrand Co., Inc., New York, 1942.

5. Buchheim, Robert W. Space Technology, Chapter 7, John Wiley and Sons, Inc., New York, 1959.

"Page missing from available version"

PAGES 12 - 13

It is seen that a transfer central angle of 1.15 rad is near optimum; however, angles between 0.7 and 1.75 rad result in less than 1 percent increase in fuel mass.

Figure 3 shows the total fuel mass versus elevation angle at initiation of main braking. The sharp rise in fuel usage with larger angles results from the increased altitude at braking initiation. For comparison, the fuel required for constant thrust descent from 185.2 km, as determined from a calculus of variations descent, for a 32,659 kg vehicle and 106,757 N thrust is approximately 20 percent greater than is required for a descent profile consisting of a 1.05 rad transfer-ellipse central angle from 185.2 km to 103 km and constant thrust from 103 km to touchdown.

DISPLACEMENT SCHEME

This scheme is simply a positional one in which the vehicle position as a function of range from the beacon is compared with a standard value. Vehicle position is established by measuring the angle between the local vertical at ignition and the radar line of sight to the beacon. The guidance equations for thrust and attitude control are

$$F_c = F_o + K_t (\dot{R}_m - \dot{R}_d)$$

and

$$\chi_c = \chi_o + bt + K_\gamma (\gamma_d - \gamma_m)$$

where F_c and χ_c are the commanded values of thrust and thrust direction. The nominal thrust level, F_o, and the nominal thrust direction program, $\chi_o + bt$, establish a standard trajectory. The value of F_o used was 80 percent of maximum thrust to allow sufficient throttle range for control. The $\chi_o + bt$ attitude program was empirically derived and almost exactly duplicates a calculus of variations descent to less than 20 km slant range. The remaining terms in the guidance equations are corrections based on deviations in range rate and position from the corresponding values from the standard trajectory. The desired values of $\dot{R}_d$ and γ_d versus range are stored in the guidance computer. The block diagram for the basic attitude system is shown in Figure 4.

This scheme could be expected to be somewhat inefficient since any displacement from standard shortly after ignition could result in large values of thrust normal to the velocity vector to displace the vehicle laterally to the desired position.

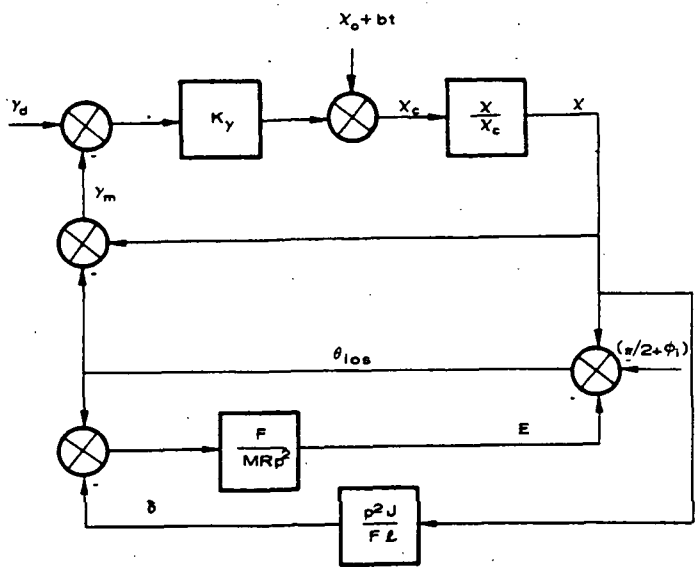


FIGURE 4. DISPLACEMENT SCHEME ATTITUDE

If the ΔV penalty were not too excessive, this disadvantage could be tolerable because of the simplicity of the scheme.

To perform a linear analysis of the system's dynamic response, certain assumptions may be made so that the system of Figure 4 reduces to that of Figure 5:

1. Roots caused by engine gimbal deflection, δ, are neglected. Analog computer runs show the almost negligible effect of this simplification for the nominal run with no initial errors.

2. A fixed point on the trajectory is used so that the nominal attitude angle, $\chi_o + bt$, is a constant value. This constant, along with the constant value of $(\pi/2 + \varphi_i)$ which arises from the problem geometry, then represents initial conditions and is omitted from the dynamic analysis.

3. A constant value of thrust is used. Since thrust variations couple into the attitude control loop as variations in the location of the open loop roots in

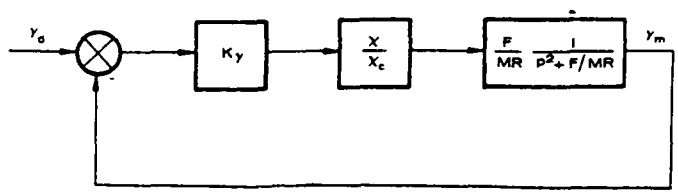


FIGURE 5. SIMPLIFIED DISPLACEMENT SCHEME ATTITUDE SYSTEM

the complex plane, a well-behaved system operating at near nominal conditions will be only slightly affected by thrust variations. The fixed thrust assumption is therefore valid for initial attitude stability studies.

The final reduced system has a closed loop transfer function denominator

$$1 + K_\gamma \left(\frac{\chi}{\chi_c}\right) \frac{F/MR}{(p^2 + F/MR)} .$$

Regardless of the form for $\left(\frac{\chi}{\chi_c}\right)$, the presence of the two undamped roots renders stabilization difficult to achieve. The required damping can be most simply accomplished by utilizing a lead network in the forward transfer function. However, the location of the compensation network poles and zeros to provide satisfactory damping of the closed loop roots requires excessively large values for the network components. An essentially similar method would utilize a rate gyro mounted on the radar antenna. The output of this gyro would be compared with another nominal curve which would have to be stored in the guidance computer. Here again however, large gain is required to produce the required damping. These high gains exaggerate any noise present in the system and, in the case of the gyro, increase errors caused by offset and measurement errors in the gyro itself.

Figure 6 shows the cross coupling effect of thrust control on the unstabilized system. These curves are the result of digital computer runs. Both curves represent nominal initial conditions and an attitude control gain of 10. Curve A is for constant thrust whereas curve B represents variable thrust. Thrust control has only a slight effect on the instability.

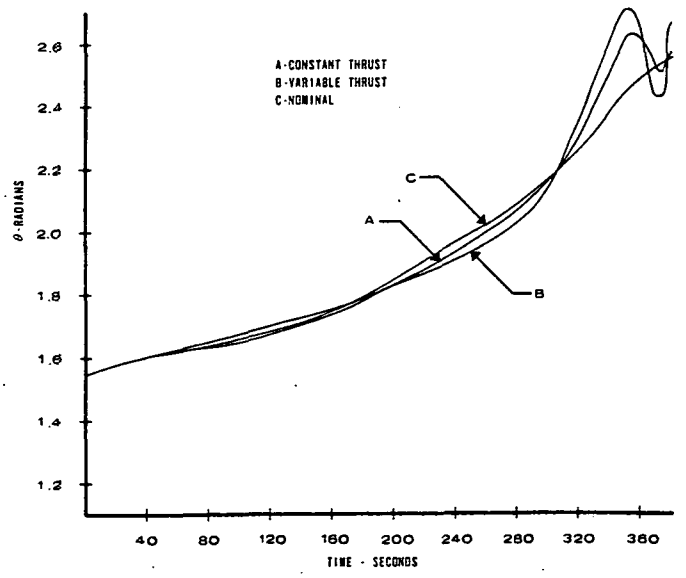


FIGURE 6. PATH ANGLE VERSUS TIME - UNCOMPENSATED DISPLACEMENT SCHEME

LINE-OF-SIGHT RATE SCHEME

The line-of-sight rate scheme is based on a programed angle between the velocity vector and line of sight to the beacon, i.e., a programed angular rate of the line of sight. The instability of the displacement scheme was not encountered; here, the nominal trajectory is defined by an angle of attack program derived from an open loop standard trajectory. A nominal program for the angle between velocity vector and line of sight is then also defined by this standard; guidance corrections are developed by measuring deviations in this angle. The guidance equations for thrust and attitude are, respectively,

$$F_c = F_o + K_t (\dot{R}_m - \dot{R}_d)$$

and

$$\alpha_c = \alpha_n + K_\gamma (\gamma_{los_n} - \gamma_{los_m})$$

where γ_{los_m} is computed from radar outputs as $\tan^{-1} R\dot{E}/\dot{R}$ or as $R\dot{E}/\dot{R}$ using the small angle approximation.

The block diagram for this system is shown in Figure 7. The angle $(\theta_{los} - \delta)$ is small and the thrust

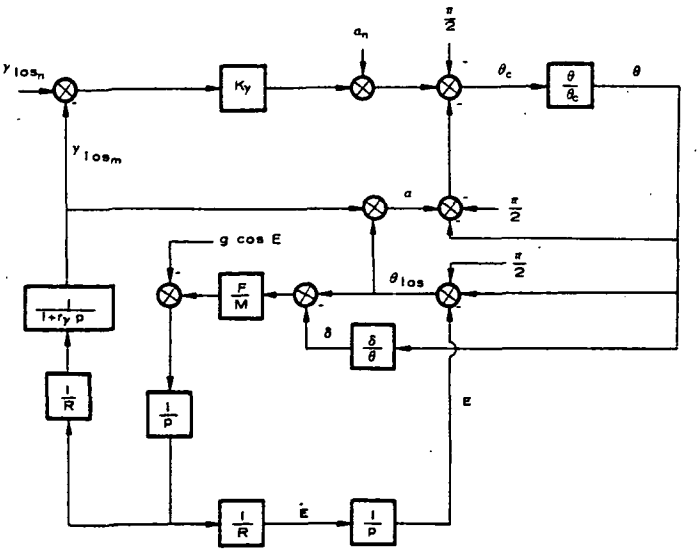


FIGURE 7. LINE-OF-SIGHT RATE SCHEME ATTITUDE

acceleration normal to the line of sight is approximated by F/M ($\theta_{los} - \delta$). The term involving τ_γ represents radar measurement time lag. As in the previous system, a constant value of thrust is used. For purposes of analysis, Figure 7 may be shown to be reducible to the form shown in Figure 8. The assumptions made in this reduction are:

1. A fixed trajectory point is assumed so that the nominal angle of attack, α_n, is a constant and, along with the various $\pi/2$ constants which arise from the geometry, is omitted from the dynamic analysis.

2. Since the lunar central angle traversed during main braking is small, the gravity acceleration normal to the line of sight is calculated from a flat surface approximation. This term is relatively constant and is neglected in the fixed point dynamic analysis.

3. Variations in the angle E are small compared with variations in the angle θ so that the value for E may be neglected and then $\theta_{los} = \theta$. This omission was shown to have negligible effect on the roots of the characteristic equation of the closed loop transfer function.

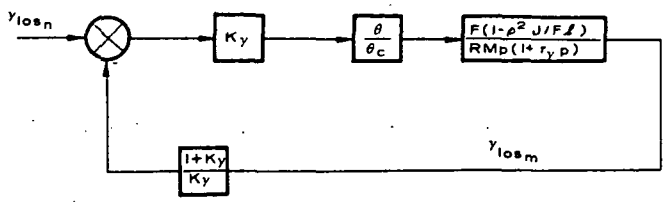


FIGURE 8. LINE-OF-SIGHT RATE SCHEME SIMPLIFIED ATTITUDE

The transfer function for the vehicle attitude control loop (Fig. 9) is

$$\frac{\theta}{\theta_c} = \frac{1 + \tau_f p}{1 + pA\tau_f + p^2 \frac{J}{K_5 F\ell} + p^3 \frac{J\tau_f}{K_5 F\ell}}$$

where the engine gimbal servo lag τ_g is assumed to be zero.

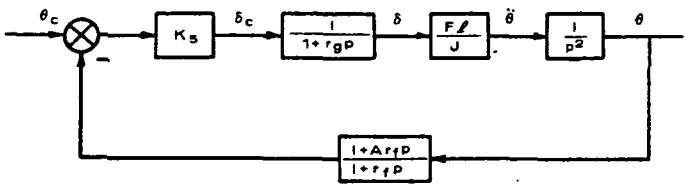


FIGURE 9. VEHICLE ATTITUDE CONTROL

The denominator of the closed loop transfer function for the system of Figure 8 is then

$$1 + \frac{F(1 + K_\gamma)\ (1 + p\tau_f)\ (1 - p^2 \frac{J}{F\ell})}{\dot{R}\ M\ (p^3 \frac{J\tau_L}{K_5 F\ell} + p^2 \frac{J}{K_5 F\ell} + pA\tau_f + 1)\ (p^2\tau_\gamma + p)}$$

and the characteristic equation is

$$p^5 + p^4 \frac{J}{K_5 F\ell}\ (\tau_\gamma + \tau_f) + p^3 \left[A\tau_f\tau_\gamma + \frac{J}{K_5 F\ell} - \frac{J\tau_f\ (1 + K_\gamma)}{\dot{R}\ M\ell} \right] + p^2 \left[A\tau_f + \tau_\gamma - \frac{J(1 + K_\gamma)}{\dot{R}\ M\ell} \right] + p \left[1 + \frac{\tau_f F}{\dot{R}M}\ (1 + K_\gamma) \right] + \frac{F\ (1 + K_\gamma)}{\dot{R}M} .$$

A representative set of coefficients for this equation is listed in Table I with the corresponding roots. These indicate well-behaved response.

TABLE I

K_5	=	.35 rad/rad			τ_γ	=	1 s		
τ_f	=	.22 s			$\dot{R}$	=	91.5 m/s		
F	=	106,757 N			R	=	1,220 m		
K	=	4 rad/rad							
	R_1		R_2	I_2		R_3		I_3	
	-1.3		-1.97	±0.626		-0.15		±0.382	

CONCLUSIONS

A transfer-ellipse central coast angle of approximately 1.15 rad was most desirable in accomplishing transition from orbit to the point at which main braking is initiated.

Two schemes for guidance during the main braking phase were compared for fuel penalty and for end point deviations at a range of 20 km. The comparisons were made for various initial conditions and measurement

errors. The displacement scheme, inherently unstable, was stabilized by comparing the output of an antenna-mounted rate gyro to a stored nominal curve of line-of-sight angular rate versus range to go. Of primary interest were deviations in velocity, elevation angle, and fuel consumption. These deviations resulted from variations about nominal in braking initiation altitude, initial central angle, thrust magnitude uncertainty, and bias error in the line-of-sight angular rate measurement.

The significant results of the comparison may be summarized as follows:

1. Both schemes showed considerably more sensitivity to bias error in measuring line-of-sight rate than to other errors.

2. The rate scheme proved more effective in handling initial altitude, central angle, and thrust uncertainty errors.

3. With the rate scheme, elevation angle errors caused by altitude and central angle errors are offset to a degree by path angle errors.

4. For both schemes, velocity at 20 km range showed only slight sensitivity to initial condition and thrust errors.

5. With the exception of angular rate bias error, the rate scheme provides less fuel mass increase per unit error than the displacement scheme.

6. While the two methods compared closely for nominal errors, the displacement scheme was unable to cope with larger errors such as initial altitude error in the low direction greater than 7 km and initial central angle error greater than 17 mrad.

Other less significant errors for both schemes included a 2 percent error and time lags up to three seconds in measuring R and $\dot{R}$, deviations in initial velocity of ± 50 m/s, and deviations in initial path angle of ± 3.75 mrad.

In view of the comparison, the line-of-sight rate scheme appears to be preferable to the scheme based on positional deviation from a standard trajectory.

C. TERMINAL GUIDANCE TECHNIQUES FOR SOFT LUNAR LANDING by George W. Neal and John F. Pavlick

A preliminary analysis of terminal guidance and control techniques necessary for landing the Saturn V LLS Configuration A vehicle on the lunar surface from altitudes of up to 20 km has been conducted. An analog computer was used to simulate the system. Since an unprepared landing site was postulated, the vehicle must be equipped with sensors to determine a suitable landing site and sufficient maneuverability to land at the selected point.

This study is concerned primarily with the maximum maneuver envelope and a desirable formulation of the guidance equations.

INTRODUCTION

An investigation of the guidance and control techniques necessary to permit the soft landing of an unmanned vehicle upon the lunar surface has been completed. Various philosophies and approaches have been presented concerning the vernier phase, i.e., the final portion of flight prior to actual contact with the lunar surface. However, it is generally conceded that a vehicle should have the ability to maneuver on its final approach to the moon to avoid touchdown in an area of large protuberances or excavations. This maneuverability should be as large as is economically feasible, considering the penalty imposed in fuel requirements and the general character of the terrain within the proposed landing area. The vehicle must be equipped with the sensors necessary to insure selection of an adequate landing site. It is extremely important that such a vehicle, while performing a terrain-avoidance maneuver, should exhibit inherently good control stability.

A method of using television information for landing site selection in the direct descent mode is described in Reference 1. The present study is an extension of this method to the case where the vehicle descends from lunar orbit. Television data are transmitted to earth where a suitable landing site is chosen. The site location is transmitted to the vehicle in terms of distance from the nominal impact point. The maneuver is then executed automatically by the vehicle. The sequence may be repeated for fine corrections as altitude is reduced. The magnitude of state variable off-nominal variation which will allow soft landing decreases rapidly below 20 km. The scope of this study was limited to determining the maximum maneuverability of the vehicle, the fuel penalties associated with lateral maneuvers, and a desirable formulation of the guidance equations.

With these major requirements in mind, several control equations were written to allow the vehicle to deviate from a nominal incoming trajectory and to maneuver until it is above a favorable landing site. The vehicle would then descend vertically, assuming that no further maneuvers were necessary, to an altitude of several meters at which time the rate of descent would be reduced to less than 5 m/s. The lateral velocity would be 1 m/s or less, the attitude angle would be $\pm$ 0.035 rad, and the vehicle angular rate would be no more than 0.009 rad/s. Upon satisfying these requirements, thrust would be terminated and the vehicle would be allowed to free fall to the lunar surface. A hover condition was not imposed since its fuel penalty is large, and a hover requirement appears to be unnecessary for unmanned vehicles.

Using the Saturn V LLS Configuration A, various maneuver trajectories were run until the control gain factors were close to optimum. Although gain optimization studies were not complete, the guidance gain factors used in this study are believed to be realistic. Upon completion of the maximum landing footprint determination, representative runs were selected and the vernier-phase initial altitude was varied throughout a range of approximately $\pm$ 20 m/s. Ideally, the thrust obtained from the engines would be equal to the thrust dictated by the throttle position. Since such an ideal situation is not likely to exist, runs were made with the actual thrust obtained equal to as little as 0.95 of the thrust commanded. Other runs were conducted which considered increases in thrust obtained to as much as 1.05 of the commanded value.

The physical characteristics of the Saturn V LLS Configuration A were obtained from Reference 2. The payload capability was about 12,700 kg when utilizing two RL-10 engines with a maximum thrust of 67,200 N each and approximately 11,560 kg of usable propellants. A sketch of the vehicle with pertinent dimensions is shown in Figure 1.

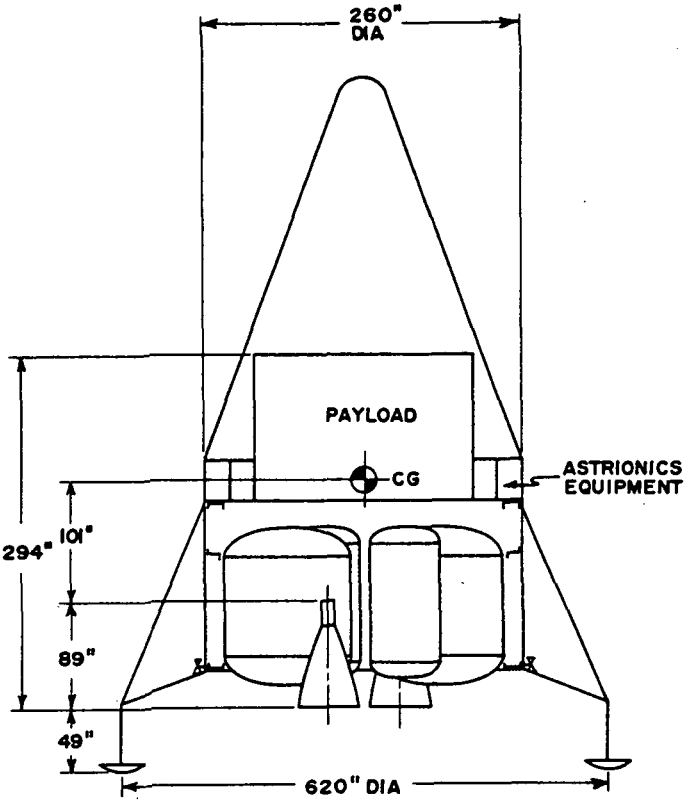


FIGURE 1. SATURN V LLS CONFIGURATION A

NOMENCLATURE

X_R	displacement of desired impact point from vernier phase initiation point, m
X_M	displacement of desired impact point from nominal, m
X	displacement along the horizontal reference axis, m
X_N	displacement along the horizontal reference axis of the nominal trajectory, m
Z	displacement along the vertical reference axis, m
θ	vehicle attitude angle (positive clockwise), rad
T_C	20 percent of maximum engine thrust, N
T	engine thrust, N
T_N	80 percent of maximum engine thrust, N
β_C	commanded value of engine gimbal angle, rad
θ_C	commanded value of vehicle attitude angle, rad
Z_{REF}	reference altitude, m

Subscripts

1	associated with engine 1
2	associated with engine 2

Dots over symbols indicate derivatives with respect to time.

ANALOG SIMULATION

The three-degree-of-freedom stability equations used in the vernier phase analog simulation were written with respect to a space-fixed, right-hand coordinate system, which is shown in Figure 2 with positive directions as indicated. The vehicle force and moment equations and the gimbal dynamics equations may be easily derived and are not presented.

System Control Equations

The commanded vehicle attitude angle control equation is a function of the maneuver required and the distance traveled (equation 1). A range rate term is used to provide damping.

$$\theta_C = K_4 (X_R - X) - K_3 \dot{X} \qquad (1)$$

$$X_R = X_N + X_M \qquad (2)$$

The term X_R is defined as the nominal range plus the desired maneuver distance and is always measured positive with respect to the vernier phase initiation point. Since the statistical probability of finding a suitable landing site within a given landing footprint is

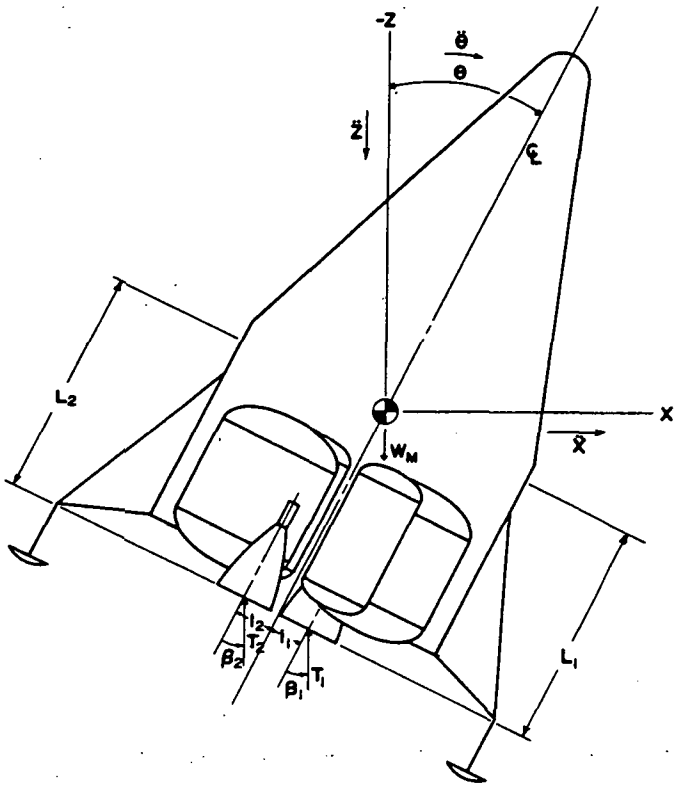


FIGURE 2. COORDINATE SYSTEM

the same both uprange and downrange, it was decided to limit this study only to downrange maneuvers to take advantage of fuel economies.

Values of the control gain factor K_4 are presented in Figure 3 as a function of the required maneuver X_M for vernier phase initiation altitudes of 10 and 20 km. The factor K_4 and the maneuver distance will be constant for any given run. However, should a new landing site be selected and the maneuver requirements

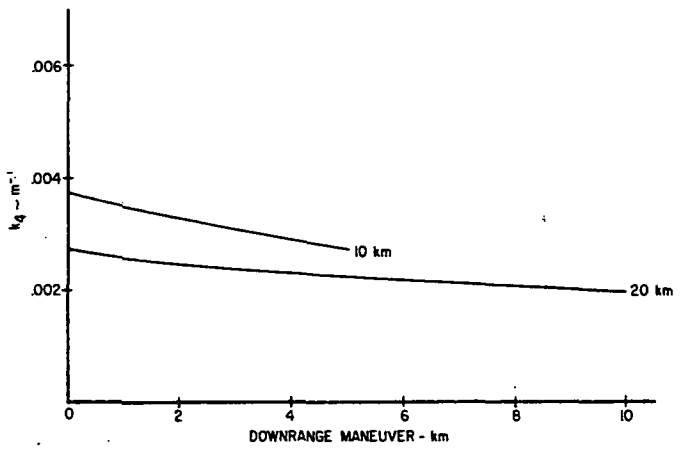


FIGURE 3. GUIDANCE GAIN FACTOR VERSUS MANEUVER FOR VERNIER INITIATION ALTITUDES OF 10 AND 20 km

be changed, a new value of K_4 must be obtained from the curves presented in Figure 3.

The commanded gimbal angle equation is a function of the commanded attitude angle θ_C as shown in equation 3 and contains an attitude angular rate term to provide the proper damping. The control gain factors K_5 and K_6 have constant values.

$$\beta_C = K_5 (\theta_C - \theta) - K_6 \dot{\theta} \quad (3)$$

The thrust control equations for both engines are defined as shown in equations 4 and 5.

$$T_1 = T_{N1} + K_9 \left[X_R - X\right] + K_1 \left[\frac{\dot{Z}}{Z} - \left(\frac{\dot{Z}}{Z}\right)_{REF}\right] + T_{C1} \quad (4)$$

$$T_2 = T_{N2} + K_{10} \left[X_R - X\right] + K_2 \left[\frac{\dot{Z}}{Z} - \left(\frac{\dot{Z}}{Z}\right)_{REF}\right] + T_{C2} \quad (5)$$

The terms T_{N1} and T_{N2} are constants and, for the purpose of this study, were assumed to be 80 percent of the maximum available thrust. A throttling ratio of 8 to 1 that allowed each engine a thrust range of 8,400 to 67,200 N was assumed. The terms T_{C1} and T_{C2} are included in the equations whenever 100 m/s $\leq \dot{Z} \leq$ 200 m/s and are equivalent to 20 percent of the maximum thrust. T_{C1} and T_{C2} allow the use of maximum thrust for the vehicle to reach the desired reference $\frac{\dot{Z}}{Z}$ relatively early in the sequence. The second term in each of the thrust equations is dependent upon the absolute value of $(X_R - X)$ and would make no contribution to the required thrust once the maneuver had been performed. The value of the reference term $\left(\frac{\dot{Z}}{Z}\right)_{REF}$ was consistent with the deceleration capability of the vehicle. The term K $\left[\frac{\dot{Z}}{Z} - \left(\frac{\dot{Z}}{Z}\right)_{REF}\right]$ was not included in the thrust equation until $\frac{\dot{Z}}{Z} \geq \left(\frac{\dot{Z}}{Z}\right)_{REF}$ or until $\dot{Z} \geq 100$ m/s. When either of these conditions was satisfied, the third term in the thrust equations was included and the necessary thrust was commanded to enable the vehicle to fly the required reference.

Study Limitations or Constraints

1. The commanded attitude angle (θ_C) was limited to 0.785 rad.

2. The engine gimbal angular rates $\dot{\beta}_1$ and $\dot{\beta}_2$ were limited to maximum values of ± 0.14 rad/s.

3. The engine gimbal angles β_1 and β_2 were limited to maximum values of ± 0.13 rad/s.

4. The thrust of each engine was limited to a maximum value of 67,200 N and to a minimum value of 8,400 N.

5. The mass of the vehicle was considered constant at 18,480 kg which compares with a nominal initial mass of 19,580 kg at 20 km and a touchdown mass of 16,920 kg.

DISCUSSION OF RESULTS

Maximum Maneuver Determination for Vernier Phase Initiation Occurring at an Altitude of 20 km.

The flight profiles for three possible descent trajectories are shown in Figure 4, which presents curves for the nominal descent and maneuvers of 5 and 10 km. A maximum maneuver of 10 km is believed adequate when vernier phase initiation occurs at 20 km.

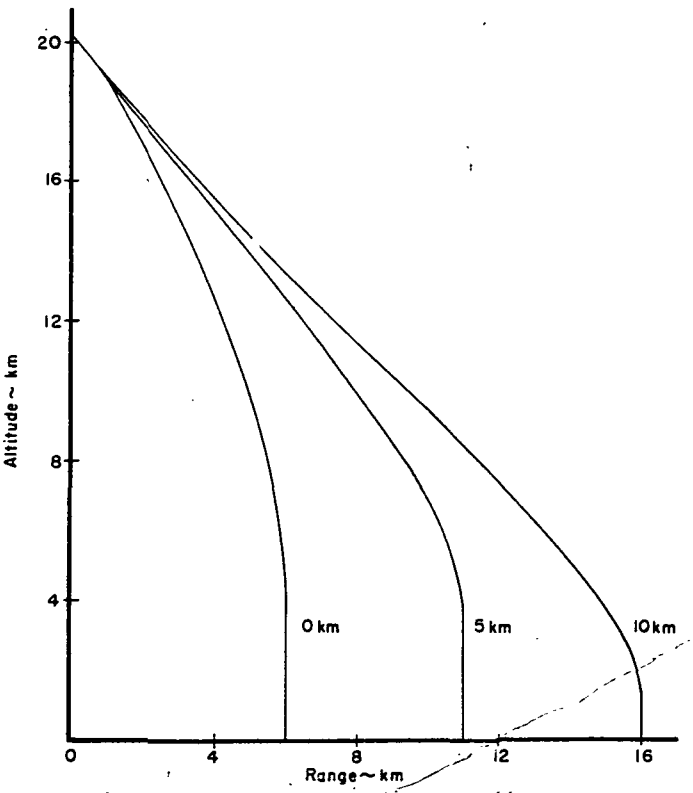


FIGURE 4. ALTITUDE VERSUS RANGE FOR MANEUVERS OF 0, 5, AND 10 km AND VERNIER INITIATION ALTITUDE OF 20 km

Downrange (positive) maneuvers impose smaller ΔV penalties than uprange maneuvers of the same magnitude. Figure 5 presents a curve of ΔV versus maneuver distance and clearly shows the excessive penalties imposed by uprange maneuvers. Range and altitude rates as well as vehicle angular rates were considered reasonable for the maneuver envelopes of this study.

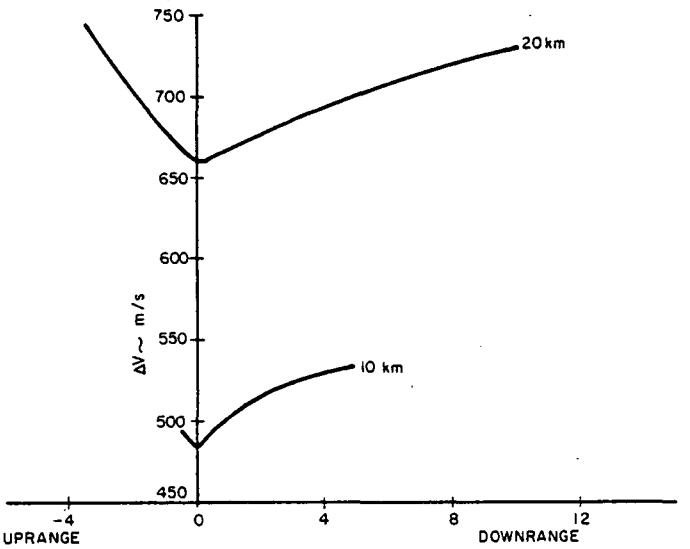


FIGURE 5. ΔV VERSUS MANEUVER FOR VERNIER INITIATION ALTITUDES OF 10 AND 20 km

Thrust Variations for Vernier Phase Initiation Altitude of 20 km.

The effects of variations in thrust on vehicle performance were investigated since it is improbable that the actual engine thrust will correspond to the commanded thrust level. It was found that the engine thrust actually obtained could be ± 5 percent of the maximum or minimum commanded engine thrust without appreciably compromising the conditions at impact. The thrust variations were conducted on maneuver runs of 0, 5, and 10 km. The vertical velocities at touchdown, for runs with thrust variations, were under 1 m/s throughout the maneuver range considered, thus clearly demonstrating the ability of the thrust control system to compensate for thrust variations.

Initial Altitude Variations for Nominal Vernier Phase Initiation Altitude of 20 km.

Runs were made that considered off-nominal initial altitudes. The results of these runs showed that

decreases in initial altitude of up to 1.5 km could be tolerated by the control system while performing maneuvers within the 0 to 10 km range. Increases in initial altitude of up to 2 km were performed with no effect on the conditions at touchdown. Off-nominal initial altitude conditions had no effect on attitude angle or horizontal velocity at touchdown.

Initial Vertical Velocity Variations for Vernier Phase Initiation Altitude of 20 km.

The nominal initial range rate was 230 m/s. Variations in initial range rate of ± 20 m/s had no effect on the vertical velocity, horizontal velocity, or attitude angle at touchdown. The data obtained indicate a limitation of ± 10 m/s in order not to exceed a nominal touchdown point variation of 100 m.

Initial Range Rate Variations for Vernier Phase Initiation Altitude of 20 km.

The nominal initial descent rate of 380 m/s was varied, and it was determined that the system was capable of compensating for increases in initial vertical velocity of up to 15 m/s while performing maneuvers of 0 to 10 km. Decreases in $\dot{Z}_i$ of up to 25 m/s resulted in a change in the shape of the thrust curve but had no effect on the conditions at impact. Variations in the initial rate of descent had no effect on the horizontal velocity or altitude angle at impact.

Maximum Maneuver Determination for Vernier Phase Initiation at an Altitude of 10 km.

The nominal (no maneuver) descent profile is shown in Figure 6 along with trajectories demonstrating the ability of the vehicle to complete downrange maneuvers of 2 and 4 km. A maneuverability of 4 km in the downrange direction covers an area that should prove sufficiently large to allow for the selection of an adequate landing site. The ΔV penalties associated with maneuvers such as those considered are shown in Figure 5 as functions of the desired maneuver.

Variations in Initial Conditions for Nominal Initiation Altitude of 10 km.

The nominal initial vertical and horizontal velocities were 269 m/s and 156 m/s, respectively. The guidance and control system of the vehicle has successfully demonstrated the capability to compensate

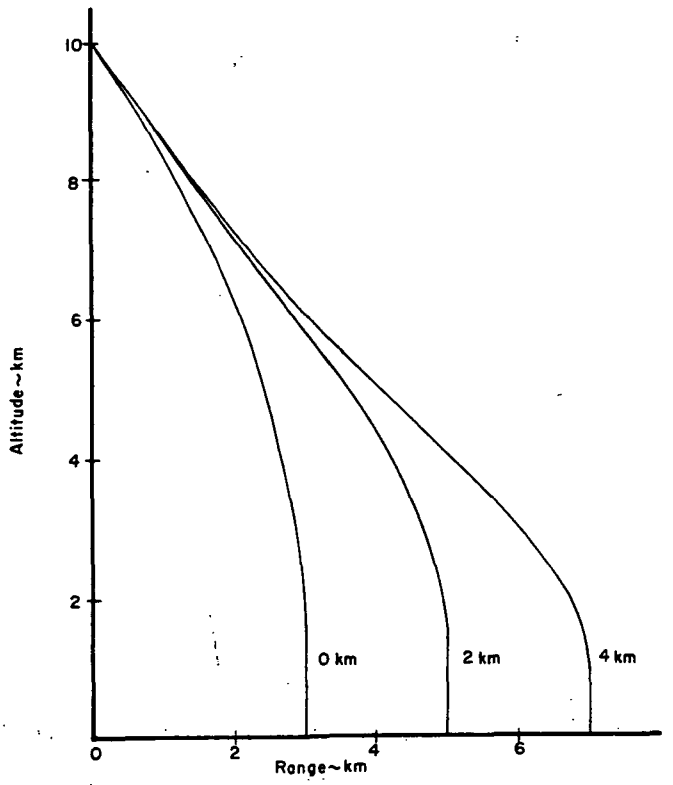


FIGURE 6. ALTITUDE VERSUS RANGE FOR MANEUVERS OF 0, 2, AND 4 km WITH VERNIER INITIATION ALTITUDE OF 10 km

for decreases in initial altitude of up to 125 m and still meet the desired conditions at impact. Increases in initial altitude of up to 500 m had no adverse effect on vehicle performance. Runs were also made that considered increases in altitude rate of 1.5 m/s and decreases of 10 m/s without compromising the desired touchdown conditions.

Variations in range rate of ± 15 m/s were accomplished with no appreciable effect on landing conditions.

Changes in the initial altitude, altitude rate, and range rate of the magnitude considered had no noticeable effect on the attitude angle, attitude angular rate, or horizontal velocity at touchdown.

Thrust Variations for Vernier Phase Initiation Altitude of 10 km.

Results of runs considering thrust variations show that the actual thrust obtained may be no more than ± 0.5 percent of the commanded value of thrust without jeopardizing the vehicle as it lands. The thrust variations considered for maneuver runs covering a range of 0 to 4 km were successfully completed.

CONCLUSIONS

The results of this investigation demonstrated the ability of the vehicle to satisfy large maneuver requirements for missions of this type. When vernier-phase initiation occurred at an altitude of 20 km, the guidance and control equations were able to compensate for large off-nominal initial conditions while delivering the vehicle to the selected landing site according to the imposed touchdown requirements.

The maneuverability of the vehicle with vernier-phase initiation occurring at 10 km is considered adequate. However, only small deviations from the nominal initial values of altitude rate and thrust could be tolerated without exceeding the desired touchdown conditions. The off-nominal initial conditions may be increased only by severely reducing the maneuver requirements imposed on the vehicle.

On the basis of these results, landing site selection for the vehicle should occur at altitudes of 10 to 20 km to allow reasonable deviations from nominal initial conditions and to maintain a satisfactory maneuver envelope.

The fuel penalty for this large maneuverability is approximately 10 percent of the fuel used in a nominal descent from the initial points used in the study and does not appear to be excessive if the maneuver is required for a soft landing on an acceptable site.

The results of this study and those in progress concerning radar and television sensor performance, lunar surface characteristics, and control system characteristics will be used to define a technically feasible and desirable system to accomplish an automatic soft lunar landing.

REFERENCES

1. MTP-M-63-1, Volume V, "Lunar Logistic System, Astrionics Aspects," March 15, 1963.

2. Memo No. M-P&VE-FS-306, "Preliminary Weights and Centers of Gravity of the Saturn V Lunar Logistics System Configuration A," March 6, 1963.

D. SPACE VEHICLE ATTITUDE CONTROL USING PSEUDO RATE MODULATION by David N. Schultz

The results of an analysis of "pseudo rate" modulation for use with a space vehicle attitude control system are presented. An on-off control system using engines with a fixed thrust level and direction is assumed to provide vehicle pitch, yaw, and roll stabilization. With a minor modification in this basic system, the concept of pseudo rate modulation can be incorporated, which simultaneously pulse-rate and pulse-width modulates the control engines.

An extensive analysis was performed to mathematically describe this concept with respect to minimum impulse limit cycle operation as well as control system stability regardless of whether an active rate sensor (rate gyro, etc.) is used. Both steady-state and transient performance of this concept are analyzed, and the results are presented in both mathematical and graphical form.

INTRODUCTION

Use of on-off pulse modulation techniques in space vehicle attitude control applications has been extensively studied and many workable schemes have been advocated (Refs. 1, 2, and 3). This article presents the results of an analysis of one particular modulation scheme called "derived rate increment stabilization" (Ref. 1) or, as commonly used in Saturn V/S-IVB studies "pseudo rate modulation."

Since the dynamic range of an attitude control system usually must be large, pulse modulation is used to fulfill the requirement that the control engines be capable of producing thrust pulses of either short or long on-times. This corresponds to minimum impulse generation and compensation of large disturbances, respectively. Over a certain range of the input signal applied to the pseudo rate modulator, the output is caused to be both pulse-width and pulse-rate modulated as a function of the input signal level. As will be shown, pseudo rate modulation has desirable characteristics of both of these modulation schemes.

To provide this type of operation, each actuator or on-off contactor* has a preset amount of deadspace, a specified amount of positive switching hysteresis, and a negative feedback consisting of a first order lag network. The deadspace is adjusted to satisfy a required amount of control system attitude deadband or attitude rate deadband (if rate control modes are used). Both the amount of switching hysteresis and the feedback network characteristics are usually sized to insure control system minimum impulse operation of the particular engines used.

An additional characteristic of the feedback lag network is the provision of a small amount of control system damping. Although this capability is sometimes not the dominant consideration for selecting the network and gain factors, it is nevertheless desirable and an important consideration for the analysis since the amount of damping is normally sufficient to maintain control system limit cycle operation in the event of a rate sensor failure. All modulator gain factors and network characteristics are mathematically analyzed to obtain all possible control system operating ranges, and the results are presented graphically.

Figure 1 shows a simplified single axis control system block diagram which includes a pseudo rate modulator. As illustrated, there are three cases for sensed rate input into the ϵ summing junction. For Case I, ϵ is equivalent to sensed attitude information only. This case will be analyzed first and is given not

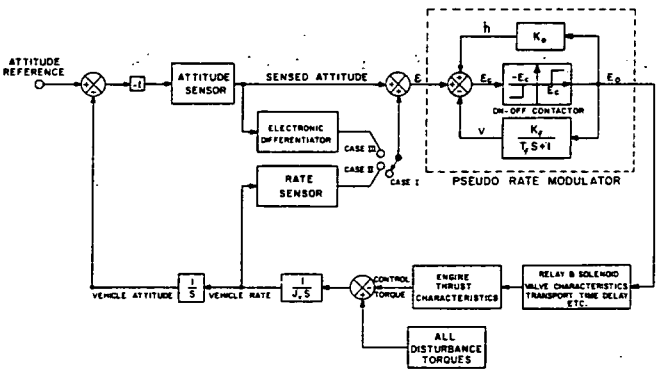


FIGURE 1. SIMPLIFIED SINGLE AXIS CONTROL SYSTEM

* This terminology and most of the nomenclature used were mutually established by R-ASTR-NGA and members of the thrust vector control study group (Ref. 4).

only to demonstrate the basic characteristics of pseudo rate modulation but also to illustrate the provision of control system damping by the lag network of the pseudo rate modulator as previously mentioned. For Case II, sensed rate information is provided using an active rate sensor; for Case III, electronic differentiation is assumed to provide the sensed rate signal. If the attitude reference is held constant with time, Cases II and III give similar single axis closed loop response. For simplicity, the analysis for these cases will be considered similar, especially if the time lags and gain factors through the respective rate sensing loops can be assumed to be the same.

Note that Figure 1 also includes the effects of engine thrust characteristics as well as relay (if used) and solenoid valve characteristics, transport time delay, etc. These effects as well as sensor lags and contactor switching hysteresis usually dictate minimum impulse operation of an on-off control system. If some sort of pulse modulation scheme were not used, large values for minimum impulse could result.

STATIC CHARACTERISTICS OF PSEUDO RATE MODULATION (INPUT HELD CONSTANT)

Regardless of the state of the system, and in terms of the parameters defined by Figure 1, the input voltage $\epsilon_c(t)$ to the on-off contactor may be expressed as

$$\epsilon_c(t) = \epsilon(t) + h(t) - V(t) \tag{1}$$

where $\epsilon(t)$ is the combined error signal voltage (sensed attitude only for Case I), $h(t)$ is the positive switching hysteresis, and $V(t)$ is the voltage of the negative feedback lag network. For $|\epsilon_c| < |E_c|$, the contactor is off; for $|\epsilon_c| \geq |E_c|$, the contactor is on with the polarity depending upon the polarity of ϵ_c.

To derive the static characteristics of the pseudo rate modulator, assume the modulator input error signal ϵ is adjusted to some constant value equal to ϵ_o. Since no steady-state pulsing action will take place for $|\epsilon_o| < |E_c|$, consider that $|\epsilon_o| \geq |E_c|$. At the instant $t = o^-$ (just prior to contactor actuation), $h = o$ and $\epsilon_c = E_c$ so that equation 1 reduces to

$$\epsilon_c = \epsilon_o - V_o = E_c \tag{2}$$

or $$V_o = \epsilon_o - E_c$$

where V_o is the initial condition of V at $t = o$. Now at the instant $t = o^+$ (just after contactor actuation), and in terms of the parameters given by Figure 1, $V(t)$ and h may be expressed as

$$V(t) = K_f E_o - (K_f E_o - V_o) e^{-t/\tau_f} \tag{3}$$

$$h = K_o E_o.$$

The contactor will remain energized until $t = t_1$ at which time $\epsilon_c(t_1) = E_c$. Assuming the instant just prior to cutoff and using equation 3, equation 1 may be written as

$$\epsilon_c(t_1) = \epsilon_o + K_o E_o - [K_f E_o - (K_f E_o - V_o) e^{-t_1/\tau_f}] = E_c. \tag{4}$$

Using V_o as given by equation 2, equation 4 may be solved for t_1 to give

$$t_1 = \tau_f \ln \left[\frac{K_f - \left(\frac{\epsilon_o - E_c}{E_o} \right)}{K_f - K_o - \left(\frac{\epsilon_o - E_c}{E_o} \right)} \right] \quad \text{where } |\epsilon_o| \geq |E_c| \quad \epsilon_o = \text{constant}. \tag{5}$$

In general, for steady-state operation, this expresses the contactor on-time increment t_1 as a function of constant values of ϵ_o and the circuit parameters K_o, K_f, τ_f, E_c, and E_o. It is interesting to solve equation 5 for $\epsilon_o = E_c$ to give the expression for the minimum on-time $t_1 = \Delta t$. This gives

$$\Delta t = \tau_f \ln \left[\frac{K_f}{K_f - K_o} \right] = \tau_f \ln \left[\frac{1}{1 - \frac{K_o}{K_f}} \right] \quad \text{expression for minimum on-time } (\epsilon_o = E_c) \tag{6}$$

$$\Delta t \simeq \tau_f \left(\frac{K_o}{K_f} \right) \quad \text{for} \quad \frac{K_o}{K_f} \ll 1.$$

The simplified expression for Δt is an approximation obtained by expanding the natural logarithm in a Maclaurin's series and considering K_o/K_f small with respect to unity. The error is less than 5 percent for values of $K_o/K_f \leq 0.1$.

Two observations may be made from the expression for minimum on-time given by equation 6. First, since for most attitude control systems it is reasonable to assume that ϵ is very nearly constant during minimum impulse limit cycle operation (Case I of Fig. 1), this equation establishes the minimum impulse of the control system. Also observe that if there is a sufficient amount of transport time delay through the engine and valve characteristics along with time lags through the sensors, equation 6 also enforces minimum impulse for Cases II and III. Second, note that equation 6 is independent of output voltage E_o (as long as E_o is not zero). This means that even if the output voltage falls off because of power supply trouble or relay or solenoid loading, the minimum impulse of the control system will not change, assuming E_c remains constant (this threshold is usually carefully controlled) and the passive network parameters K_o, K_f, and τ_f do not change.

Returning now to the general expression for on-time (equation 5), observe that

$$t_1 \to \infty \text{ as } \frac{\epsilon_o - E_c}{E_o} \to (K_f - K_o). \tag{7}$$

Since relations 7 show the contactor to be continuously on for this condition of input, the modulation range may be defined as

$$0 \leq \frac{\epsilon_o - E_c}{E_o} \leq (K_f - K_o). \tag{8}$$

This range may be normalized to

$$0 \leq \Psi \leq 1 \tag{9}$$

where

$$\Psi \equiv \frac{\epsilon_o - E_c}{E_o\,(K_f - K_o)} = \text{normalized input.}$$

In terms of normalized input, equation 5 may be rewritten

$$t_1 = \tau_f \ln \frac{K_f - \Psi\,(K_f - K_o)}{(K_f - K_o)\,(1 - \Psi)} \quad 0 \leq \Psi \leq 1 \tag{10}$$

as the general expression for on-time increment as a function of the normalized input Ψ.

Similar reasoning may be used to derive the expression for the off-time increment as a function of constant values of modulator input error signal ϵ_o. Since only time increments are of interest, the origin of the time axis may be redefined such that at the instant just before contactor cutoff, $t = 0^-$. Equation 1 then gives

$$\epsilon_c = \epsilon_o + h - V_1 = E_c \tag{11}$$

or

$$V_1 = \epsilon_o - E_c + K_o E_o$$

where V_1 is the initial condition of $V(t)$ at this instant of time. Just after cutoff, $t = 0^+$, and using equation 11, $V(t)$ may be expressed as

$$V(t) = V_1 e^{-t/\tau_f} = (\epsilon_o - E_c + K_o E_o)\, e^{-t/\tau_f}. \tag{12}$$

Considering $t = t_2$ to be the instant just prior to the time the contactor is again actuated and using equations 1 and 12,

$$\epsilon_c(t_2) = \epsilon_o - V(t_2) = E_c$$

$$\epsilon_o - (\epsilon_o - E_c + K_o E_o) e^{-t_2/\tau_f} = E_c. \tag{13}$$

Solving equation 13 for the off-time increment t_2 gives

$$t_2 = \tau_f \ln \left[\frac{\dfrac{\epsilon_o - E_c}{E_o} + K_o}{\dfrac{\epsilon_o - E_c}{E_o}} \right]. \tag{14}$$

Again using the normalized input as given by expressions 9 allows equation 14 to be rewritten as

$$t_2 = \tau_f \ln \frac{\Psi(K_f - K_o) + K_o}{\Psi(K_f - K_o)} \quad 0 \leq \Psi \leq 1 \tag{15}$$

which is the general expression for the off-time increment. Observation of equation 15 reveals that

$$t_2 \to \infty \text{ as } \Psi \to 0. \tag{16}$$

This illustrates the important characteristic of pseudo rate modulation that, for minimum impulse operation, the off-time increment approaches a high value (theoretically infinity) as the minimum on-time Δt is accurately timed to a fixed specified amount. This long off-duration usually gives the control system time to recover from all lags around the various loops.

To graphically portray the on- and off-time increments versus the normalized input Ψ, equation 6 will be used to eliminate τ_f from equations 10 and 15 to give

$$\frac{t_1}{\Delta_t} = \frac{\ln\left[\dfrac{1-\Psi\left(1-\dfrac{K_o}{K_f}\right)}{\left(1-\dfrac{K_o}{K_f}\right)(1-\Psi)}\right]}{\ln\left[\dfrac{1}{1-\dfrac{K_o}{K_f}}\right]} \quad \text{normalized on-time increment} \qquad (17a)$$

$$\frac{t_2}{\Delta t} = \frac{\ln\left[\dfrac{\Psi\left(1-\dfrac{K_o}{K_f}\right)+\dfrac{K_o}{K_f}}{\Psi\left(1-\dfrac{K_o}{K_f}\right)}\right]}{\ln\left[\dfrac{1}{1-\dfrac{K_o}{K_f}}\right]} \quad \text{normalized off-time increment} \qquad (17b)$$

$$0 \le \left[\Psi = \frac{\epsilon_o - E_c}{K_f E_o - K_o E_o}\right] \le 1 \quad \text{normalized input.} \qquad (17c)$$

Figure 2 graphs expressions 17 which (more fundamentally) normalize the on- and off-time increments against Δt which is usually specified. The normalized increments $t_1/\Delta t$ and $t_2/\Delta t$ are graphed as a function of Ψ with K_o/K_f as a varied parameter. The value of K_o/K_f must be different from zero but is usually as small as possible (for reasons given later) so the

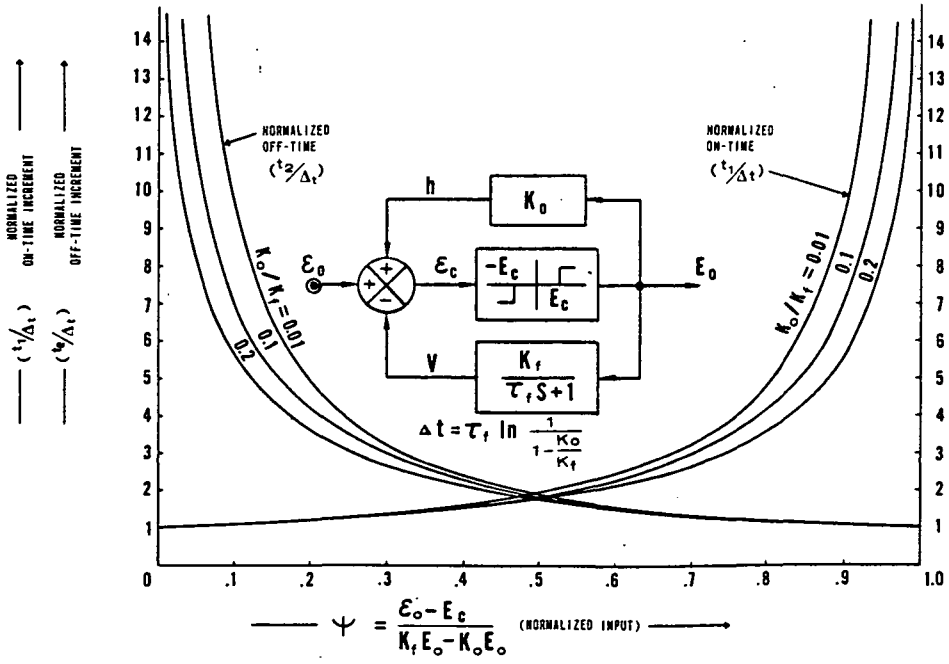


FIGURE 2. NORMALIZED ON-AND OFF-TIME INCREMENTS VERSUS NORMALIZED INPUT

maximum value for this ratio considered in Figure 2 is 0.2. With reference to Figure 2, as Ψ increases from 0 to 0.5, $t_1/\Delta t$ (on-time) slowly increases while $t_2/\Delta t$ (off-time) rapidly decreases. Just the reverse occurs for Ψ increasing from 0.5 to 1.0 since $t_1/\Delta t$ rapidly increases toward the full-on condition while $t_2/\Delta t$ decreases slowly toward the fixed level of 1.0. This is characteristic of a combination of pulse width and pulse rate modulation thereby enabling desirable characteristics of both modulation schemes to be obtained. These are the previously mentioned long off-time intervals (pulse rate modulation) during minimum impulse limit cycle operation and the pulse widening effect (pulse width modulation) to give maximum restoring torque capability when needed. Also observe from Figure 2 that, for a given value of K_o/K_f, $\Psi = 0.5$ when $t_1 = t_2$. This is easily verified by solving expressions 17 for Ψ when $t_1 = t_2$.

A deeper insight into the nature of pseudo rate modulation may be obtained by using expressions 17 to plot $t_1/\Delta t$ versus $t_2/\Delta t$ holding K_o/K_f constant as the normalized input Ψ is varied. This same graph also provides a convenient comparison of pseudo rate modulation with both pulse rate and pulse width modulation techniques. Referring to the insert of Figure 3, pulse rate modulation is represented as a straight horizontal line passing through $(t_1/\Delta t)_{min}$ as the normalized input Ψ is varied because the on-time must remain fixed as the off-time is varied (Ref. 2). Pulse width modulation however, retains a constant period or carrier frequency (Ref. 2) as both the on- and off-times are varied. The period may be expressed as

$$\frac{P}{\Delta t} \approx \frac{t_1}{\Delta t} + \frac{t_2}{\Delta t}. \qquad (18)$$

Now for $\frac{P}{\Delta t}$ equal to a positive constant, equation 18 represents a straight line with negative slope on the $t_1/\Delta t$ versus $t_2/\Delta t$ graph; as Ψ is varied, pulse width modulation is also easily depicted on this graph (see insert of Fig. 3). For pulse width modulation, the on-time goes to zero as Ψ becomes zero ($\epsilon_o \rightarrow E_c$); also the off-time goes to zero as Ψ increases (ϵ_o increases). In certain instances this is disadvantageous especially for very short on-times since control system power can be consumed with no generation of thrust (i.e., the pulses are too short for the solenoid valves to respond). Pseudo rate modulation may now

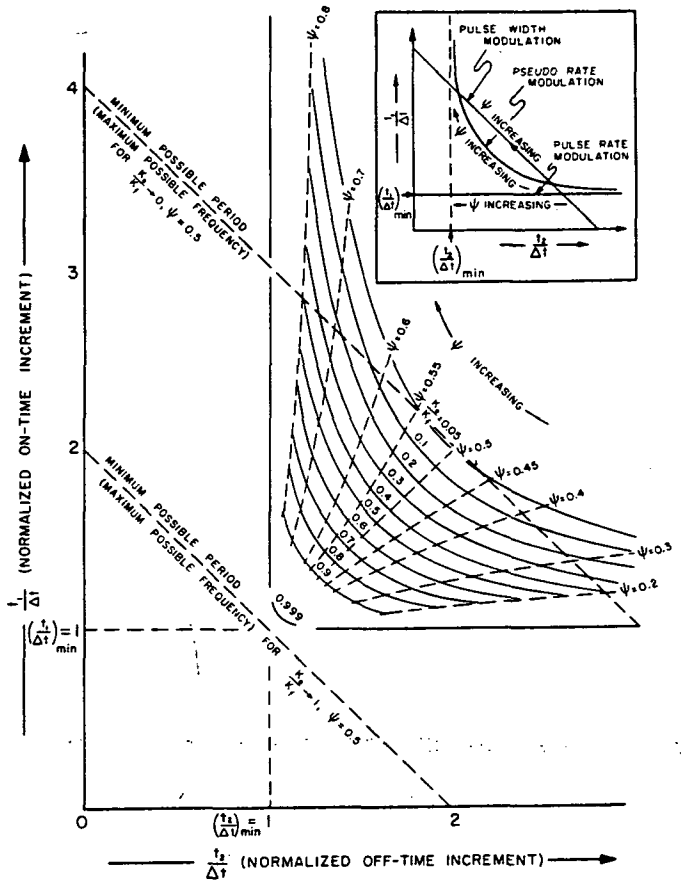


FIGURE 3. NORMALIZED ON-TIME VERSUS NORMALIZED OFF-TIME

be seen to be a combination of both. For small Ψ, all lines of constant K_o/K_f are asymptotic to $(t_1/\Delta t)_{min}$ which represents pulse rate modulation; for midrange values of Ψ ($0.4 < \Psi < 0.6$), the system very nearly displays the characteristics of pulse width modulation. As Ψ changes over all possible values and considering steady-state operation for each, neither the on-time nor the off-time can become zero (each approaches a fixed specified level). Therefore, during steady-state operation, pseudo rate modulation does not have the disadvantage of a continuously vanishing on-time that is characteristic of pulse width modulation.

An important characteristic immediately obvious from Figure 3 is that for $\Psi = 0.5$ the period of the modulator pulse waveform is always minimum for a given K_o/K_f thereby giving the maximum steady-state pulse frequency. Substituting expressions 17 into equation 18 for $\Psi = 0.5$ gives

$$\left(\frac{P_{min}}{\Delta t}\right)_{\Psi = 0.5} = 2\left[1 - \frac{\ln\left(1 + \frac{K_o}{K_f}\right)}{\ln\left(1 - \frac{K_o}{K_f}\right)}\right] \qquad (19)$$

$$\left(\frac{P_{min}}{\Delta t}\right)_{\Psi = 0.5} \approx 2\left[2 - \frac{K_o}{K_f}\right] \text{ for } \frac{K_o}{K_f} << 1.$$

This simplified expression for $\left(\frac{P_{min}}{\Delta t}\right)_{\Psi = 0.5}$ gives less than 1 percent error when $K_o/K_f \leq 0.1$ and was obtained using the first two terms of a Maclaurin's series expansion for the natural logarithm with all second order terms neglected.

With reference to equation 6, the range with K_o/K_f varied from 0 to 1 corresponds to variation of $\Delta t/\tau_f$ from 0 to ∞, thereby covering all possible variations of these parameters with respect to the minimum period (maximum frequency) of the pseudo rate modulator output waveform. Letting K_o/K_f approach 1 and 0, respectively, in equation 19 gives

$$\lim_{K_o/K_f \to 1} \left(\frac{P_{min}}{\Delta t}\right)_{\Psi = 0.5} = 2;$$

$$\lim_{K_o/K_f \to 0} \left(\frac{P_{min}}{\Delta t}\right)_{\Psi = 0.5} = 4$$

therefore,

$$2 \leq \left(\frac{P_{min}}{\Delta t}\right)_{\Psi = 0.5} \leq 4 \text{ as } 1 \geq \frac{K_o}{K_f} \geq 0. \qquad (20)$$

This range of $\left(\frac{P_{min}}{\Delta t}\right)_{\Psi = 0.5}$ applies for all possible values for K_o, K_f, and τ_f. Figure 4 graphs equation 19, and the range of $\left(\frac{P_{min}}{\Delta t}\right)_{\Psi = 0.5}$ given by relations 20 is clearly shown. To minimize the design value for maximum steady-state frequency of the particular modulator used, Figure 4 and relations 20 show that

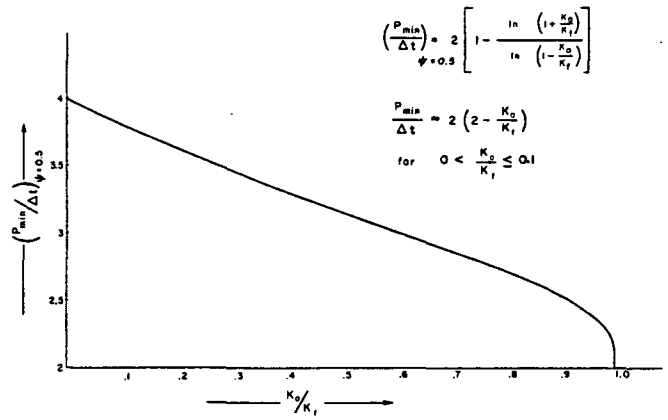

$\left(\frac{P_{min}}{\Delta t}\right)_{\Psi=0.5} = 2\left[1 - \frac{\ln\left(1+\frac{K_o}{K_f}\right)}{\ln\left(1-\frac{K_o}{K_f}\right)}\right]$
$\frac{P_{min}}{\Delta t} = 2\left(2 - \frac{K_o}{K_f}\right)$
for $0 < \frac{K_o}{K_f} \leq 0.1$
$\left(P_{min}/\Delta t\right)_{\Psi=0.5}$
K_o/K_f

FIGURE 4. GRAPH OF $\left(\frac{P_{min}}{\Delta t}\right)_{\Psi=0.5}$ VERSUS $\frac{K_o}{K_f}$

K_o/K_f must be small. For example, if $\Delta t = 0.05$ s and K_o/K_f is made as small as possible, then $P_{min} = 4\Delta t = 0.2$ s and the maximum steady-state pulse frequency of the modulator output would be 5 pulses per second. The pulse frequency would be less than 5 pulses per second for any other value of Ψ (Fig. 3).

It is of interest to plot $t_1/(t_1 + t_2)$ versus Ψ since this is the variation of average modulator output which is proportional to control torque (control system considered to be open looped). Using equations 17, Figure 5 is given to show this relationship. Even though the curves are nonlinear, note that as K_o/K_f becomes smaller, the average output waveform becomes more linear as a function of the normalized input Ψ.

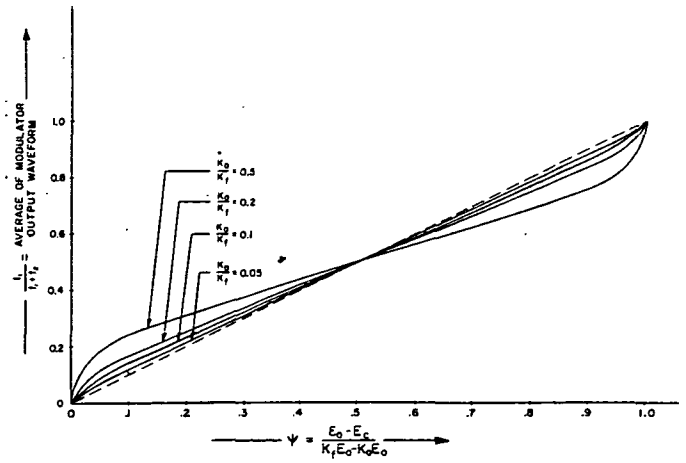

$\frac{t_1}{t_1+t_2}$ = AVERAGE OF MODULATOR OUTPUT WAVEFORM
$\frac{K_o}{K_f}$ = 0.5
$\frac{K_o}{K_f}$ = 0.2
$\frac{K_o}{K_f}$ = 0.1
$\frac{K_o}{K_f}$ = 0.05
$\Psi = \frac{E_o - E_c}{K_f E_o - K_o E_o}$

FIGURE 5. AVERAGE OF MODULATOR OUTPUT WAVEFORM VERSUS NORMALIZED INPUT

Up to this point, no restrictions have been placed on the variables defining the modulation range Ψ. However, for $|\epsilon| > |E_c|$ and if ϵ is suddenly reduced to zero (because of loss of signal or an intentional command), it is possible for the contactor to be inadvertently energized if $|V(t)| \geq |E_c|$ at this instant. This condition will be called a "reverse firing" since the polarity of the contactor output will be opposite to that of ϵ prior to loss of signal; the time duration can be considerable if the lag network has charged to its maximum voltage. The maximum value for $|V(t)|$ is $|K_f E_o|$ and the requirement for the contactor to remain off is that $|\epsilon_c| < |E_c|$. Using equation 1 with $\epsilon(t) = 0$ and the fact that $h(t)$ must go to zero when the contactor turns off would give

$$|K_f E_o| \leq |E_c| \tag{21}$$

or

$$K_f \leq \frac{E_c}{E_o} .$$

If it is assumed that noise is present at the contactor input even though inequality 21 is satisfied, the contactor can still be momentarily triggered into a reverse firing and can then be held on by $h = K_o E_o$ for a time interval as long as Δt. Therefore, the absolute requirement for no reverse firings is specified by

$$|K_f E_o| + |K_o E_o| \leq |E_c|$$

or

$$K_f\left(1 + \frac{K_o}{K_f}\right) \leq \frac{E_c}{E_o} . \qquad \text{requirement for no reverse firings} \tag{22}$$

Thus,

$$(K_f)_{max} = \frac{E_c}{E_o\left(1 + \frac{K_o}{K_f}\right)} .$$

As Ψ varies from 0 to 1, the modulation range that satisfies the requirement of no reverse firings may be obtained by substituting inequality 22 into expression 9 to give

$$0 \leq \frac{\epsilon_o - E_c}{E_o} \leq \left(\frac{E_c}{E_o} - 2K_o\right). \tag{23}$$

This latter expression indicates that for values of $|\epsilon| \geq |2E_c| - |2K_o E_o|$, the contactor must be continuously on if reverse firings are to be avoided.

As previously mentioned, it is usually desirable to keep the K_o/K_f ratio small, which requires that K_o should be made small and/or K_f should be made large. Usually, there is some lower limit for the value of K_o as dictated by the hardware used; however, the bottom equation of expressions 22 specifies an upper limit for the value of $K_f = (K_f)_{max}$. Use of $(K_f)_{max}$ then gives the smallest possible value for K_o/K_f if it is required to avoid reverse firings.

DYNAMIC CHARACTERISTICS OF PSEUDO RATE MODULATION (INPUT TIME VARYING)

Even though equation 1 describes the input voltage to the on-off contactor in general, up to this point it has been used with constant values ($\epsilon = \epsilon_o$) for modulator input error voltage to derive the steady-state or static characteristics of pseudo rate modulation. These static characteristics describe closed loop on- and off-times over any time interval such that $\epsilon(t)$ can be assumed to be very nearly constant (during Case I minimum impulse operation, for example).

It is now of interest to analyze the dynamic characteristics by allowing $\epsilon(t)$ to vary as a function of the control system sensed inputs. Neglecting all disturbance torques and assuming idealized sensor characteristics, Figure 1 is redrawn as Figure 6. For Cases II and III and in terms of the variables defined in this figure, $\epsilon(t)$ may be expressed as

$$\epsilon(t) = a_0\phi(t) + a_1\dot{\phi}(t) \tag{24}$$

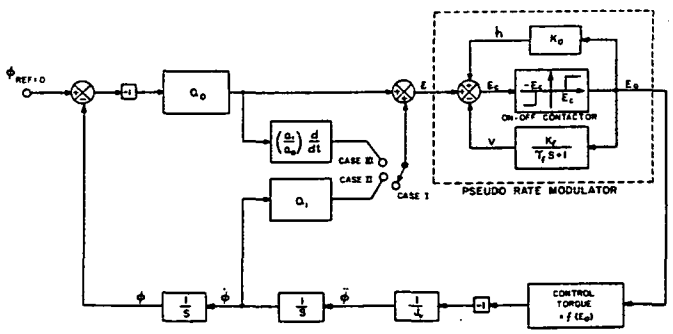


FIGURE 6. CONTROL SYSTEM WITH IDEALIZED SENSOR CHARACTERISTICS

where $\phi(t)$ is the vehicle attitude for $\phi_{REF} = 0$, $\dot{\phi}(t)$ is the vehicle angular rate, and a_o and a_1 are the respective control gain factors. By setting $a_1 = 0$, equation 24 then applies for Case I. If it is now assumed that the vehicle angular acceleration $\ddot{\phi}$ is constant whenever control torque is generated, the vehicle attitude and rate may be written as

$$\dot{\phi}(t) = \dot{\phi}_o + \ddot{\phi}t \tag{25}$$

$$\phi(t) = \phi_o + \dot{\phi}_o t + \frac{1}{2}\ddot{\phi}t^2$$

where ϕ_o and $\dot{\phi}_o$ are any specified set of initial conditions of attitude and rate at $t = 0$. Substitution of equation 25 into equation 24 gives

$$\epsilon(t) = \epsilon_o + \dot{\epsilon}_o t + \frac{1}{2}\ddot{\epsilon}_o t^2 \qquad \text{Cases II and III} \quad \ddot{\phi} = \text{constant} \quad (\text{for Case I, } a_1 = 0) \tag{26}$$

where

$$\epsilon_o = a_0\phi_o + a_1\dot{\phi}_o$$

$$\dot{\epsilon}_o = a_0\dot{\phi}_o + a_1\ddot{\phi}$$

$$\ddot{\epsilon}_o = a_0\ddot{\phi}.$$

This equation describes the input voltage to the pseudo rate modulator as a function of time and the initial state (ϕ_o and $\dot{\phi}_o$) of the vehicle with time measured from the instant that constant control acceleration is applied. Whenever $\ddot{\phi} = 0$, $\ddot{\epsilon}_o$ is also zero so equation 26 reduces to

$$\epsilon(t) = \epsilon_o + \dot{\epsilon}_o t \qquad \text{Cases II and III} \quad \ddot{\phi} = 0 \quad (\text{for Case I, } a_1 = 0) \tag{27}$$

where

$$\epsilon_o = a_0\phi_o + a_1\dot{\phi}_o$$

$$\dot{\epsilon}_o = a_0\dot{\phi}_o.$$

Equation 27 again describes the input voltage to the pseudo rate modulator with the instant of contactor cutoff now defining both the time reference origin and the initial conditions ϕ_o and $\dot{\phi}_o$.

Several interpretations may be given to equations 26 and 27. For example, the overall control system could be idealized by assuming that whenever the contactor is on, $\ddot{\phi}$ is constant (equation 26 applies) and whenever the contactor is off, $\ddot{\phi}$ is zero (equation 27 applies). Because of the previously mentioned transport time delays through the engine and solenoid valve characteristics, it could also be assumed that time intervals exist such that the contactor is on, yet no control torque is produced. Equation 27 applies during these on-time intervals. Even though various other combinations are possible, to show the damping capability of pseudo rate modulation, the idealized control system will be assumed.

Damping of an on-off control system is normally depicted on the system phase plane ($\dot{\phi}$ versus ϕ) by a specified amount of slope of the on-off switching lines. This slope is usually a function of the rate to attitude gain ratio a_1/a_o of the system. As previously mentioned, even though active rate sensing is not used (i.e., $a_1 = 0$), a system using pseudo rate modulation still possesses a certain amount of closed loop control system damping. This may be shown by deriving the equation for the cutoff boundary as illustrated by Figure 7 which is a plot of $\dot{\epsilon}$ versus ϵ instead of $\dot{\phi}$ versus ϕ since equations 26 and 27 show these to be equivalent. Since control system operation is similar for both polarities of ϵ, the cutoff boundary for positive ϵ only will be considered. The cutoff boundary is defined as

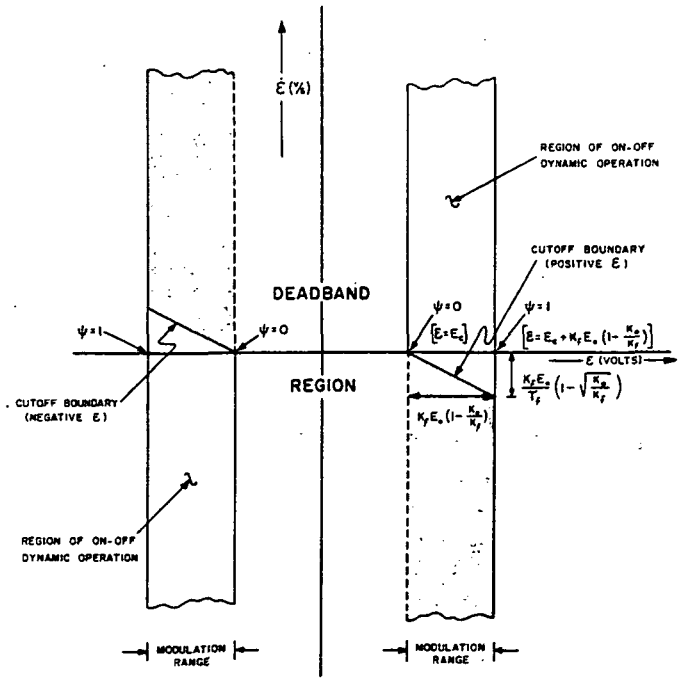


FIGURE 7. PHASE PLANE REPRESENTATION FOR PSEUDO RATE MODULATION CUTOFF BOUNDARIES (CASE I)

the locus of points ($\dot{\epsilon}_o$, ϵ_o) within the modulation range $0 \le \Psi \le 1$ such that $\dot{\epsilon}(t) = \dot{V}(t)$ when the contactor switching condition $\epsilon_c(t) = E_c$ is met. On one side of this boundary (shaded portion of Fig. 7), $|\epsilon(t)|$ decreases faster than $|V(t)|$ after cutoff. Since V(t) always tends to turn the contactor off because it is fed back negatively (with the exception of reverse firings), the contactor will remain off as the control system returns to the specified deadband. On the other side of the boundary, however, $|V(t)|$ decreases faster than $|\epsilon(t)|$, thereby causing the contactor to be reactivated before the deadband is reached.

Assuming $\ddot{\phi}$ and t are zero at the instant of contactor cutoff, $\epsilon(t)$ and V(t) for any time after cutoff are given by equations 27 and 12, respectively. The voltage rates $\dot{\epsilon}(t)$ and $\dot{V}(t)$ are given by

$$\dot{\epsilon}(t) = \dot{\epsilon}_o$$

$$\dot{V}(t) = -\frac{V_1 e^{-t/\tau_f}}{\tau_f} = -\frac{(\epsilon_o - E_c + K_o E_o}{\tau_f} e^{-t/\tau_f}. \quad (28)$$

Defining t_2 as the time just prior to a possible contactor reactivation and since h = 0, the contactor switching condition $\epsilon_c(t_2) = E_c$ gives

$$\epsilon(t_2) - E_c = V(t_2) \quad (29)$$

or

$$(\epsilon_o + \dot{\epsilon}_o t_2) - E_c = (\epsilon_o - E_c + K_o E_o) e^{-t_2/\tau_f}.$$

Setting $\dot{\epsilon}(t_2) = \dot{V}(t_2)$ in equation 28 and eliminating t_2 from both equations 28 and 29 permit the cutoff boundary to be expressed as

$$\tau_f \dot{\epsilon}_o = -[(\epsilon_o - E_c) + K_o E_o] e^{(1 + \frac{\epsilon_o - E_c}{\tau_f \dot{\epsilon}_o})} \quad (30)$$

which represents the required ($\dot{\epsilon}_o$, ϵ_o) locus as a function of the parameters τ_f, K_o, E_c, and E_o. Whenever the control torque goes to zero for the assumed idealized control system and if the initial conditions at cutoff are within the modulation range Ψ and are as specified by equation 30, there will be no more contactor actuations as the input error signal $\epsilon(t)$ drifts into the specified deadband. Equation 30 may be normalized by dividing by $K_f E_o$ to give

$$Y = -[X + \frac{K_o}{K_f}] e^{(1 + \frac{X}{Y})} \quad (31)$$

where

$$X = \frac{\epsilon_o - E_c}{K_f E_o}$$

$$Y = \frac{\tau_f \dot{\epsilon}_o}{K_f E_o}.$$

Even though equation 31 cannot be explicitly solved for Y as a function of X, a digital computer solution was obtained with Figure 8 representing the graph of the resulting data for the normalized cutoff boundary.

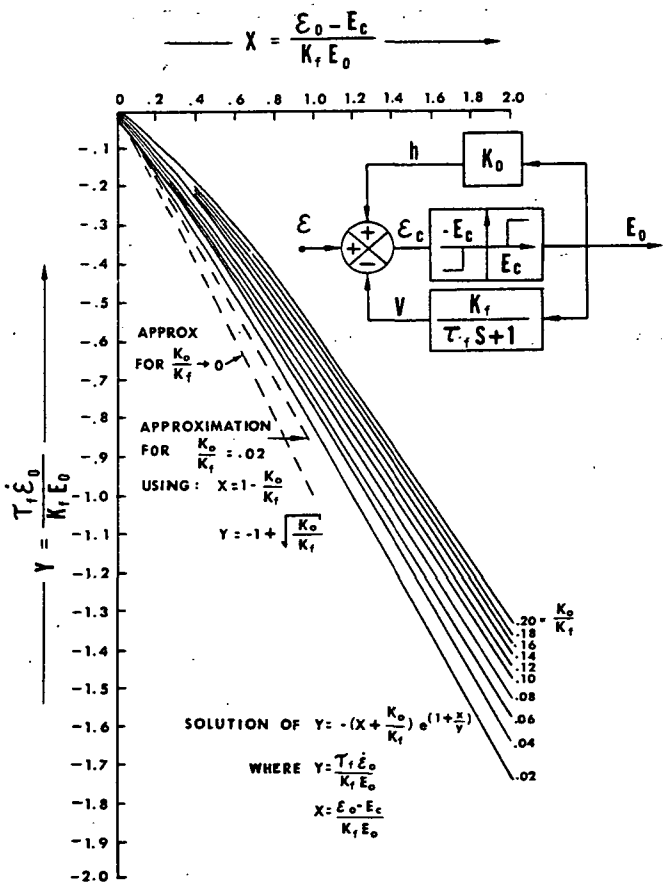

X = ε₀ - E_c / K_f E₀
0 .2 .4 .6 .8 1.0 1.2 1.4 1.6 1.8 2.0
-.1 -.2 -.3 -.4 -.5 -.6 -.7 -.8 -.9 -1.0 -1.1 -1.2 -1.3 -1.4 -1.5 -1.6 -1.7 -1.8 -1.9 -2.0
Y = τ_f ε̇₀ / K_f E₀
h
K₀
ε
ε_c
-E_c
E_c
E₀
V
K_f / τ_f S+1
APPROX FOR K₀/K_f → 0
APPROXIMATION FOR K₀/K_f = .02
USING: X = 1 - K₀/K_f
Y = -1 + √(K₀/K_f)
.20 .18 .16 .14 .12 .10 .08 .06 .04 .02 K₀/K_f
SOLUTION OF Y = -(X + K₀/K_f) e^(1+X/Y)
WHERE Y = τ_f ε̇₀ / K_f E₀
X = ε₀ - E_c / K_f E₀

FIGURE 8. NORMALIZED CUTOFF BOUNDARY

Using the definition of X from equation 31, equation 9 for $\Psi = 1$ gives

$$X = 1 - \frac{K_o}{K_f} . \tag{32}$$

Substituting this into equation 31 then yields

$$Y = -e^{(1 + \frac{1 - K_o/K_f}{Y})} \tag{33}$$

$$\lim_{K_o/K_f \to 0} Y = -e^{(1 + \frac{1}{Y})}$$

The last of these expressions has the solution $Y = -1$ and from equation 32, $X = -1$ for $(K_o/K_f) \to 0$. It would therefore seem reasonable to approximate the cutoff boundary for $(K_o/K_f) \to 0$ in the X - Y plane by a straight line passing from the origin to the point $X = -1$, $Y = -1$ as indicated by the lower dashed line of Figure 8. For K_o/K_f small, this indicates that the exponent of equation 33 is also small, allowing the exponential to be expanded in a Maclaurin's series to give

$$Y = -\left[2 + \frac{1 - \frac{K_o}{K_f}}{Y}\right] \tag{34}$$

or

$$Y^2 + 2Y + 1 - \frac{K_o}{K_f} = 0.$$

Thus,

$$Y = -1 + \sqrt{\frac{K_o}{K_f}} .$$

For K_o/K_f small, the cutoff boundary can then be approximated by a straight line in the X-Y plane passing from the origin to the point $X = 1 - \frac{K_o}{K_f}$ (equation 32), $Y = -1 + \sqrt{\frac{K_o}{K_f}}$ (equation 34).

Referring to Figure 8, the approximation for $K_o/K_f = 0.02$ is shown by the second dashed line which is reasonably close to the actual curve for this value of K_o/K_f. The estimate is conservative since the approximate cutoff line everywhere falls below the actual curve. Again using the definitions of X and Y from equation 31, equations 32 and 34 may be written as

$$\dot{\epsilon}_o = E_c + K_f E_o \left(1 - \frac{K_o}{K_f}\right) \tag{35}$$

$$-\dot{\epsilon}_o = \frac{K_f E_o}{\tau_f} \left(1 - \sqrt{\frac{K_o}{K_f}}\right)$$

which are the coordinates that specify the cutoff boundary at $\Psi = 1$ as illustrated by Figure 7. From both the approximate and exact expressions for $\dot{\epsilon}_o$ (equations 35 and 30, respectively), maximum damping is obtained for τ_f as high as possible (i.e., minimum slope of the cutoff boundary). Now with reference to equation 6 and assuming Δt to be a fixed specification, it may be observed that another reason to make K_o/K_f as small as possible would be to obtain a high value for τ_f, which corresponds to maximum damping. Figure 7 indicates, however, that the damping because of pseudo rate modulation alone (Case I) is only obtainable within the modulation range which is bounded by equation 23 if reverse firings are to be avoided.

Since the damping derived from pseudo rate modulation can be approximated as a straight cutoff line on the phase plane ($\dot{\epsilon}$ versus ϵ), the cutoff slopes will be additive over the modulation range when active rate damping is used (Cases II and III). Figure 9 is given to pictorially illustrate this effect.

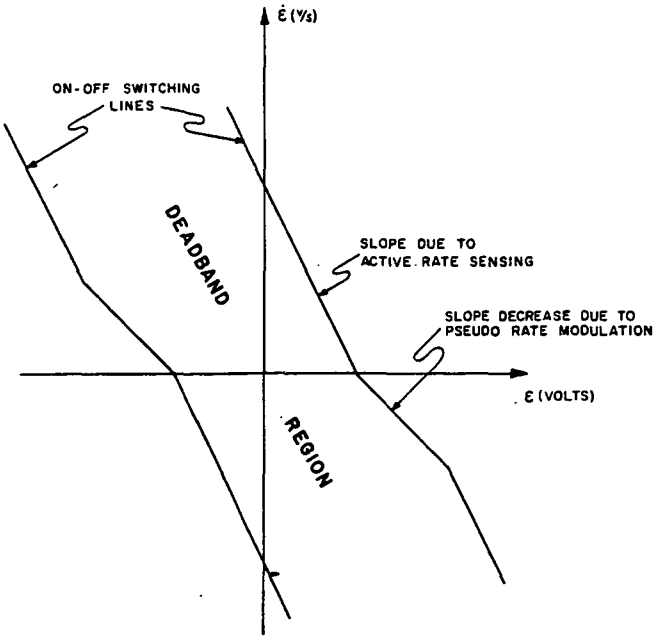


FIGURE 9. PHASE PLANE OF SYSTEM (BOTH ACTIVE RATE SENSING AND PSEUDO RATE MODULATION)

The dynamic on- and off-times within the modulation ranges illustrated by Figures 7 and 9 could also be derived using equations 26 and 27 along with reasoning similar to that used in the static characteristic analysis. However, since Reference 4 adequately covers the dynamic on-off time derivations for an idealized control system, these derivations will not be included.

CONCLUSIONS

An analysis to mathematically describe the concept of pseudo rate modulation for use with space vehicle attitude control has been performed. Both static and dynamic characteristics were investigated, depending upon the assumptions for the modulator input.

The static characteristics can be used to describe closed loop control system on- and off-times over any time intervals for which the modulator input is approximately constant. These characteristics show that the dominant feature of pseudo rate modulation is to enforce minimum impulse operation for a control system which includes appreciable effects of engine and solenoid valve characteristics, transport time delays, sensor lags, etc. As shown by the static characteristic equations and associated graphs, the off-time increment (t_2) approaches a high value (theoretically infinity) as the minimum on-time (Δt) is accurately timed to a fixed specified value. This is characteristic of pulse rate modulation and the long off-times usually give the system time to recover from all lags around the control loops during steady-state limit cycle operation. Pseudo rate modulation has the additional benefit that the minimum on-time is independent of modulator output voltage and, therefore, of output loading (as long as E_o remains different from zero). With increase of signal into the modulator, the pulse widening effect (characteristic of pulse width modulation) also gives maximum restoring torque capability when needed.

As K_o/K_f is varied from 0 to 1 (which corresponds to the variation of $\Delta t/\tau_f$ from 0 to ∞), it has been shown (expression 20 and Fig. 4) that the minimum period for the steady-state modulator output waveform varies from 4 to 2 times the specified minimum on-time Δt. Therefore, to minimize the design value for maximum steady-state frequency of the particular modulator used, the K_o/K_f ratio should be kept as small as possible. An important restriction that must be placed on the modulator gain factors and modulation range is obtained from the requirement of avoiding reverse firings. In essence, this restriction places an upper limit on the value for K_f. Use of $(K_f)_{max}$ always yields the smallest possible K_o/K_f ratio where the lower limit for the value of K_o is usually dictated by the hardware used.

The dynamic characteristics were then introduced in a general way to allow complete determination of modulator on- and off-times as a function of the state of the control system and assumptions made. For example, the damping capability of pseudo rate modulation was shown by deriving the cutoff boundary in terms of $\dot{\epsilon}$ versus ϵ. This capability can only be obtained within the modulation range (Fig. 7), and maximum damping is obtained with the value of τ_f as high as possible (equations 30 and 35). To satisfy the requirement for minimum on-time Δt, equation 6 shows that the K_o/K_f ratio should again be minimized to obtain a high value for τ_f and, therefore, maximum damping.

Assuming an idealized control system with negligible transport time delay, it is possible for on-time intervals less than the specified Δt to be produced when active rate sensing is used (Cases II and III of Fig. 6). Since the effect of the sensed rate signal is to provide control system stability, this signal will help the lag network of the pseudo rate modulator turn the contactor off. To obtain on-time intervals appreciably less than Δt, either the value for angular acceleration $\ddot{\phi}$ must be high or the rate gain factor a_1 must be large. Since active rate sensing is responsible for an early cutoff, it may be observed that it is not possible for on-times less than the specified Δt to be generated unless actual thrust of the engines is produced. Usually there is a sufficient amount of transport time delay through the engine and valve characteristics along with time lags through the sensors so that on-time intervals considerably less than Δt will not be generated.

REFERENCES

1. Nicklas, J. C., and Vivian, H. C., "Derived-Rate Increment Stabilization: Its Application to the Attitude Control Problem," Joint Automatic Control Conference, Boulder, Colorado, June 28-30, 1961, Paper No. 61-JAC-9.

2. Schultz, David N., "Space Vehicle Attitude Stabilization Using Pulse Modulated Jet Control," MTP-M-G&C-61-32, July 13, 1961.

3. Schultz, David N., "Saturn S-IVB Attitude Control Using Minimum Hardware Pulse Width Modulation," M-ASTR-NAV 461/62, September 27, 1962.

4. Thrust Vector Control Study Group, "Study of S-IVB Attitude Control System With Pseudo Rate Modulation," Auburn University, Seventh Technical Report, October 29, 1963.

E. PILOT PARTICIPATION IN SATURN V GUIDANCE AND CONTROL by Robert W. Gunderson and Gordon H. Hardy *

A piloted control system configuration was determined for the attitude stabilization and structural load relief control functions during first stage of Saturn V boost. Simulation and analytical studies assumed adequacy of perturbation equations, permitting five-degrees of rigid-body control freedom, plus the first and second modes of structural bending dynamics. Conventional control system implementation was imposed as a design requirement, with the exception of standard aircraft display panel instrumentation. A limiting factor in system performance was caused by undesirable filter response modes, introduced by flexible body stabilization filters in the rate augmentation loops.

INTRODUCTION

The advent of the manned launch vehicle has created an increasing interest in the question of pilot participation in guidance and control during boost phase of the flight profile. In particular, it has been conjectured, based primarily upon high performance aircraft experience, that overall mission reliability may be increased by inclusion of the pilot at a proper point in the guidance and control loop. Previous studies (Refs. 1, 2, 3, and 4) indicate that the question does warrant serious consideration. Other than by such general statements, however, evaluation of piloted system potential should be restricted to the framework provided by specific vehicle configurations and mission profiles.

The Astrionics Laboratory is currently engaged in a joint study with the Ames Research Center (ARC) to determine the piloted control system capabilities for the Saturn V launch vehicle and the feasibility of utilizing any such capabilities to obtain an increase in system reliability. Through the joint study approach, it has been possible to combine the booster guidance and control experience of MSFC with the piloted systems research experience of ARC.

The results discussed in this article were obtained during the introductory phase of the investigation and formed the basis for an initial report on manual booster control, provided to the NASA Research Advisory Committee on Guidance and Control, Research and Technology, dated October 3, 1963. The objective of the introductory study was primarily to determine the basic control system configuration required to permit piloted control system accomplishment of the attitude stabilization and structural load relief control functions. No attempt was made to obtain an "optimal" configuration. The intent was to determine (1) an acceptable configuration, utilizing conventional control system implementation and standard aircraft display instrumentation, and (2) those areas requiring design refinement, which should be emphasized to obtain increased system performance.

It was assumed that the system configuration would be essentially determined by the requirements imposed during the region of boost flight corresponding to maximum dynamic pressure. Consequently, the simulation studies were concentrated on a fixed time of flight corresponding to high-q, with fixed time of flight checkpoints shortly after first stage liftoff and immediately before first stage cutoff. The adequacy of this assumption was later verified during the continuous time simulation of the "performance evaluation" studies.** It should also be emphasized that the introductory studies were conducted on a fixed-base assumption; i.e., neglecting the dynamic flight environment of the Apollo capsule. The validity of fixed-base simulation was based upon both the results of the previously referenced studies and the piloted systems experience of ARC. However, the validity will be further examined through moving-base simulation on the ARC centrifuge, which was initiated in February 1964.

The authors gratefully acknowledge the simulation and system development assistance of Charles O. Jones, MSFC, and James V. West, ARC. The efforts and comments of pilot participants from the Ames Research Center, Flight Research Center, and Astronaut Office of Manned Spacecraft Center are also gratefully acknowledged.

*Research Scientist, Ames Research Center

** The performance evaluation studies have since been completed and will be reported in a joint ARC-MSFC technical report to be issued in 1964.

NOMENCLATURE

α	angle of attack
β_c	engine deflection command signal
β	engine deflection
ϕ	attitude angle
$\dot{\phi}$	attitude rate
$\dot{\phi}_f$	filtered attitude rate signal
$\Delta\phi$	attitude error
$\ddot{\gamma}$	body-fixed lateral acceleration
δ	controller output signal
δ_f	filtered controller signal
a_1	rate gain
M_x	rigid body bending moment at station x
$\frac{\partial M_x}{\partial \alpha}$	function of mass distribution and aerodynamic loading, constant at a given time of flight at station x
$\frac{\partial M_x}{\partial \beta}$	function of mass distribution and thrust, constant for a given time of flight at station x
p	Laplace operator

DESIGN CONSTRAINTS

As previously mentioned, a principal constraint placed on the control system development was the insistence on subsystems already accepted as flight hardware. Only electromechanical and hydraulic components of the conventional control system were assumed, with the exception of standard aircraft display instrumentation. Trajectory data and other design information were obtained from the Saturn V Launch Vehicle Design Data Book (Ref. 5).

In this introductory study, emphasis was placed on performance measures referring to structural loading constraints, since trajectory control during S-IC stage is considered a less stringent constraint under current design philosophy. The rigid body bending moment was computed during each simulation run according to the expression

$$M_x = \frac{\partial M_x}{\partial \alpha} + \frac{\partial M_\beta}{\partial \beta}$$

at a predetermined critical structure location. The computed moments in the pitch and yaw axis were normalized to unity at design loads and squared, summed, and recorded. The square root of the sum was then taken manually for the data presentation. Data necessary for computation of the moment resulting from bending mode acceleration terms were not available at the time of the introductory study. Consequently, a design goal of minimizing bending accelerations was imposed, but no measure for defining the degree of success was available except a comparison with conventional control system results. This was done by continuously recording the generalized bending acceleration terms.

The final performance measure used was the Cooper Pilot Opinion Scale, shown in Figure 1. This rating is the pilot's subjective opinion of task difficulty and system performance.

	ADJECTIVE RATING	NUMERICAL RATING	DESCRIPTION	PRIMARY MISSION ACCOMPLISHED?
NORMAL OPERATION	Satisfactory	1	Excellent, includes optimum	Yes
		2	Good, pleasant to fly	Yes
		3	Satisfactory, but with some mildly unpleasant characteristics	Yes
EMERGENCY OPERATION	Unsatisfactory	4	Acceptable, but with unpleasant characteristics	Yes
		5	Unacceptable for normal operation	Doubtful
		6	Acceptable for emergency condition only *	Doubtful
NO OPERATION	Unacceptable	7	Unacceptable even for emergency condition *	No
		8	Unacceptable - dangerous	No
		9	Unacceptable - uncontrollable	No
	Unprintable	10	x!@~~#~~!! Did not get back to report!	what mission?

*(Failure of a stability augmenter)

FIGURE 1. RATING SCALE

SIMULATION

Perturbation equations describing the rigid body motion were assumed throughout the introductory simulation and, as used, permitted control tasks involving five degrees of control freedom. The symmetrical configuration of the Saturn V vehicle permits aerodynamic

and inertia product cross-coupling to be neglected; however, nonlinearities of the perturbation equations were retained. The effect of these terms was also found to be negligible. In addition to rigid body dynamics, flexible body equations of motion were included up to the second mode. Hydraulic system dynamics were obtained by programing the differential equations corresponding to a p-plane curve fit of experimentally determined gain-phase plots. Stabilization filters were simulated according to their mathematical models.

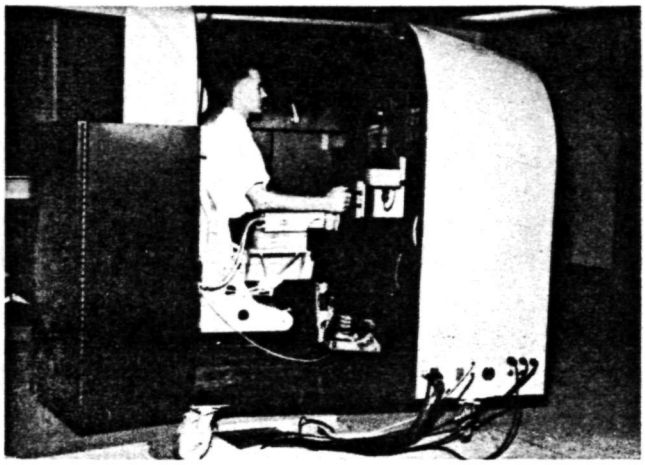


FIGURE 2. FIXED-BASE COCKPIT

Figure 2 is a photograph of the fixed-base cockpit showing the three-axis side-arm controller mounted on the arm of the chair. Also shown are the rudder pedals, used early in the simulation with a two-axis pencil-stick controller. Figure 3 is a detailed photograph of the display panel. The two large instruments

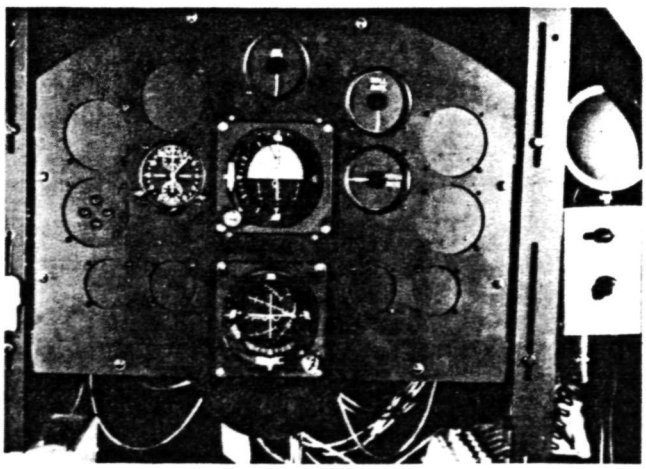


FIGURE 3. PRESENT DISPLAY PANEL

are standard aircraft all-axis attitude indicators; the upper one was used as the primary flight instrument. The display configuration presents vehicle attitude on the sphere of the upper indicator, while the body-fixed lateral accelerations were presented on the flight-director needles. Pitch and yaw attitude errors were presented on the flight-director needles of the lower instrument; however, the sphere was not used. Vehicle attitude rates were presented on the dc meters at the upper center and upper right of the display. A standard aircraft elapsed-time clock is mounted at the upper left.

Changes to the instrument panel for future studies will include:

a. Additional instrumentation for "situation" variables, such as altitude and velocity.

b. A compact, integrated display to replace the lower all-axis attitude indicator.

c. A modified all-axis attitude indicator with pitch on the inner gimbal to replace the upper all-axis attitude indicator.

The primary disturbance assumed for these studies corresponded to a horizontal wind increasing from 0 to 75 m/s at a slope of 10 m/s^2. The wind direction was rotated before each run but direction was constant during the run. The pilot was not given wind direction information prior to run initiation. In addition to the steady-state wind, a turbulence condition was supplied by a random noise generator and superimposed on the basic wind profile, with a maximum peak-to-peak amplitude of 7.5 m/s.

CONTROL SYSTEM CONFIGURATION

The block diagram in Figure 4 describes the basic pitch or yaw control system as determined by the introductory studies. The roll channel configuration is similar but does not include the lateral acceleration display loop. Attitude and attitude error signals are assumed to be supplied by the ST-124-2 stabilized platform. Rates and lateral accelerations are sensed by body-mounted rate gyros and accelerometers. Attitude stabilization was accomplished by having the pilot attempt to null attitude errors presented on the flight-director needles of the lower instrument (pitch and yaw) and on the sphere of the upper instrument (roll). Full scale attitude error presentation corresponded to ± 10 degrees. Load relief, or structural

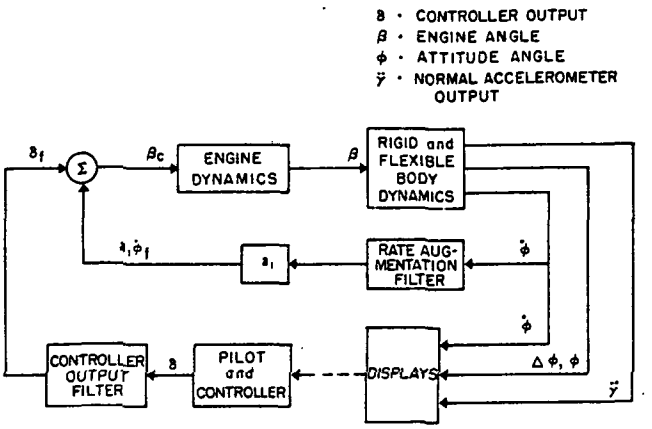
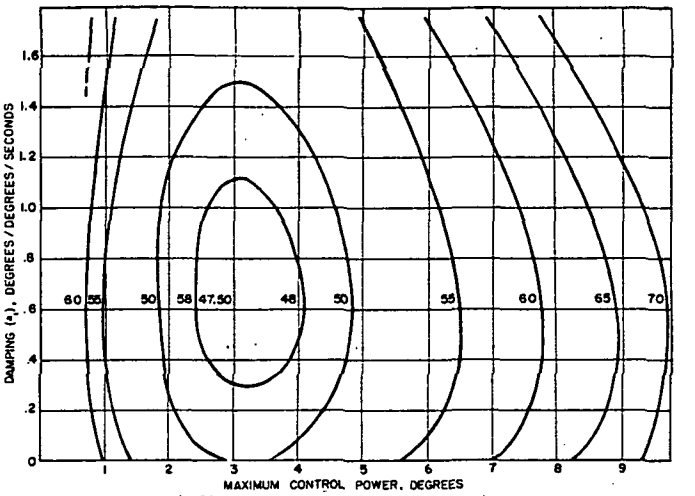


FIGURE 4. PITCH OR YAW CONTROL SYSTEM CONFIGURATION

loading control, was accomplished by having the pilot null the pitch and yaw accelerations sensed by the body-fixed accelerometers. Full scale acceleration signals corresponded to ± 0.125 g's. While accomplishing the load control task, the pilot was further instructed to null the acceleration signals, consistent with minimum engine angles and "smooth" controller motions.

RIGID BODY CONTROL

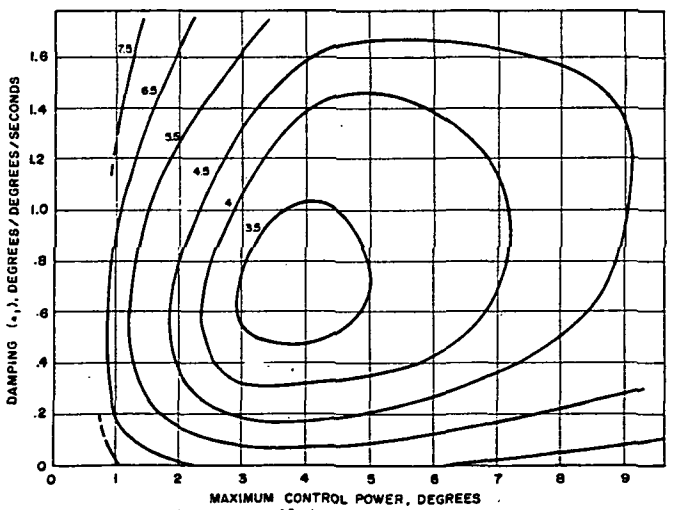
In the design approach, the basic system parameters were first determined for satisfactory accomplishment of the load control task. In addition, the pilot was required to control roll attitude, but permitted pitch and yaw attitudes to drift. Plots of control power (engine angle, β, commanded by full controller deflection) versus rate gain (a_1) were then obtained under rigid body assumptions, with pilot performance rating as a parameter in one case and bending moment loads in the other. A tentative accelerometer location was chosen at the instantaneous center of rotation corresponding to the high-q time of flight. Figure 5 shows control power versus rate gain with the root mean square of the pitch and yaw bending moments at station 3400 (normalized to unity at design loads) as a parameter. Control power and rate gain were varied identically in the pitch and yaw channels. Roll gains were kept constant at previously determined values. It can be seen from these curves that, under rigid body assumptions, the pilot could perform the three-axis control task and remain within design loads without the aid of rate augmentation. Roll attitude was controlled within ± 5 degrees error in all of the controllable cases. Pitch and yaw attitude errors would



NOTE: 1. Curves indicate percent design moment
2. Rigid body, 3 axis
3. Fixed point, high-q
4. Minimum-load control task

FIGURE 5. PILOTED BOOST PERFORMANCE

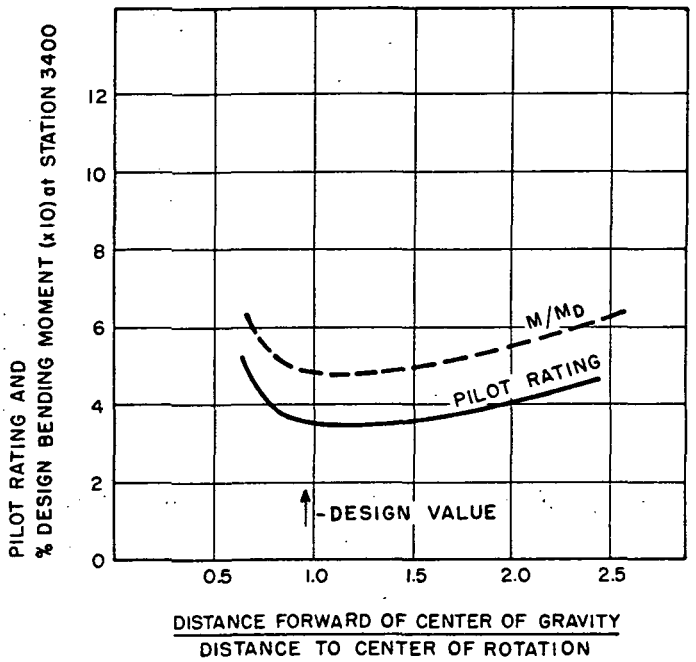
drift to error values of ± 10 degrees in about 15 to 20 seconds. Figure 6 presents pilot ratings of system handling characteristics for the same parameter variations as Figure 5. Note that the region corresponding to the most satisfactory pilot rating agrees roughly with the minimum loading region of Figure 5. Also shown is the pilot opinion that the no rate augmentation case would be acceptable only for emergency situations. Based upon these results, nominal gain values of 0.86 for rate gain and 3.0 for control power were used during the remainder of the study. After obtaining the design gains, a study was performed to



NOTE: 1. Curves indicate pilot rating
2. Rigid body, 3 axis
3. Fixed point, high-q
4. Minimum-load control task

FIGURE 6. PILOTED BOOST CONTROL BOUNDARIES

determine the sensitivity of accelerometer location for accomplishment of the load relief task. These results are shown in Figure 7. It can be seen that both pilot opinion and bending moment performance are rather insensitive to variations forward of the center of rotation, but greater sensitivity is shown for variations aft of that location.



NOTE: 1. Rigid body, 3 axis
2. Fixed point, high-q
3. Minimum-load control task
4. Rate augmentation design damping and control power

FIGURE 7. DISPLAY ACCELEROMETER LOCATION

FLEXIBLE BODY CONSIDERATIONS

From experience in the design of conventional control systems, it was anticipated that the critical design and performance problems would be created by introducing flexible body dynamics. In the case of Saturn V, the relatively small separation between rigid body control frequencies and the first bending mode natural frequency further complicates the design problem.

Figure 8A is a simplified root locus plot of an aerodynamically unstable vehicle showing system response as seen from the point of pilot insertion, shown in Figure 4. A closed rate augmentation loop is assumed but no bending dynamics stabilization filter. The two poles on the real axis represent the dominant

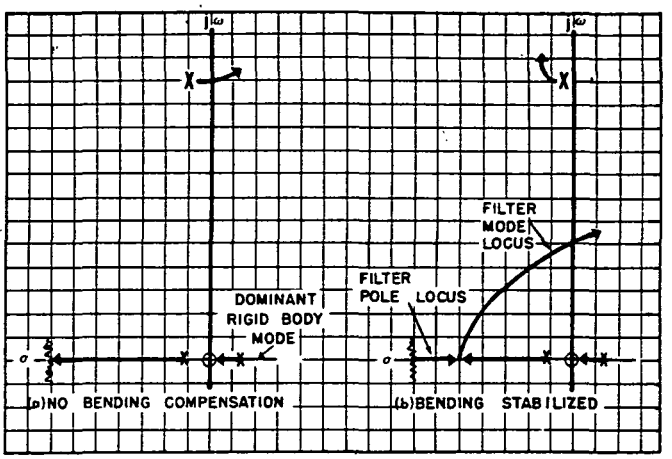


FIGURE 8. SIMPLIFIED ROOT LOCUS PLOT FOR THE INNER LOOP OF A MANUAL CONTROL SYSTEM

rigid body response, while the flexible body response is represented by the complex pole near the imaginary axis. The zero at the origin is introduced by the rate augmentation loop and, as the rate gain is increased, the roots move along the indicated loci. In particular, the locus originating at the bending pole will generally cross rapidly into the right half plane, and the system will be unstable at bending mode frequencies. To stabilize the bending dynamics, a filter is inserted into the rate loop to phase- or gain-stabilize throughout the expected ranges of parameter variation. Figure 8B illustrates the general effect of adding such compensation to the rate loop. Notice particularly that, while the locus of the divergent root is relatively unchanged, an additional locus has resulted and the dominant vehicle response is now largely determined by a combination of the divergent mode and the newly introduced filter mode. Thus, caution must be exercised in providing bending dynamics stabilization that does not degrade rigid body handling qualities. The pitch and yaw channel filters determined for the introductory control configuration are given by the transfer function

$$F(p) = \frac{336}{(p+6)(p+7)(p+8)} .$$

The filter response mode introduced by this filter was heavily damped and found to be satisfactory for nominal parameters.

In general, even though the bending dynamics may be stabilized, the closed loop bending modes will remain relatively lightly damped. It remains then to assure that:

a. flexible body motions do not obscure the rigid body control motions at the pilot's display

b. the component of bending moment caused by bending acceleration terms remains small

c. flexible body motion cues at the pilot station do not degrade pilot performance.

Control of the Saturn V during powered flight of the S-IC stage is maintained through swiveling the four outboard F-1 engines. Early in the simulation studies, it was noted that the principal source of flexible body dynamics excitation was attributable to the pilot's engine command signals. Consequently, a "smoothing" filter was placed at the output of the controller, as shown in Figure 4, to reduce the bending frequency component of the command signal. Only second order controller-output filters were investigated during this preliminary study, with the configuration described by

$$F(p) = \frac{(2.7)^2}{p^2 + 2.7p + (2.7)^2}$$

resulting in the most satisfactory effect. Lowering the filter's natural frequency caused an increase in rigid body bending moments, as a result of excessive system lag sensed by the pilot; however, raising the natural frequency provided insufficient signal attenuation at the bending mode frequencies. After determining the controller-output filter, the effect of adding an additional filter at the acceleration display was investigated. It was found, however, that the benefit gained by reducing the bending frequency component of the display signal was negated by the additional system lag. This problem was expected, since the controller and display filter lag effects are additive at the point of pilot insertion. A trade-off study, made with respect to the effects of the two filters, resulted in eliminating the display filter and retaining the original controller filter configuration.

Flexible body simulation results, using the control system configuration, indicated an approximately ten percent increase in structural loads over the rigid body case for nominal gains and coefficients. Contrary to the rigid body case, however, rate augmentation was found to be definitely required for accomplishment of the three axis task. Single axis control could be accomplished without rate augmentation, but would be recommended only for emergency situation conditions. It should be realized that these results may be improved by refinements in the control system configuration. In particular, considerable effort will be expended toward determining more optimal filter configurations. The configurations given here should by no means be considered the final or recommended solutions to the flexible body dynamics problem.

CONCLUSIONS

The following conclusions are based on the contents of this article and must be qualified by the extent of the simulation; that is, a five-degree of control freedom, fixed-based, fixed time of flight, simulation of rigid body plus first and second mode structural bending dynamics.

a. The preliminary control system configuration, obtained during the course of the introductory studies, indicates that a piloted control system is capable of completing the attitude stabilization and load relief control tasks during first stage of Saturn V boost.

b. A limiting factor in system performance was found to be attributable to the response mode introduced by the bending dynamics stabilization filters. Extensive study of this problem area is required.

c. The single-axis no-rate control tasks can be completed with the present configuration, but would be recommended only for emergency situations. Further study of the three axis case is required.

d. Structural bending motions sensed at the pilot station do not appear to be as significant a motion cue problem as originally anticipated. However, investigation of the effect will be continued into the moving-base simulation.

e. The presently used three-axis side-arm controller permits satisfactory pilot performance. Conventional aircraft instrumentation satisfied display panel requirements.

REFERENCES

1. Boeing Company, "A Pilot in the Booster Control Loop Study," USAF Contract 33(657)-7132, Boeing Document D2-80762, Final Report, Vol. I, December 1962. (Confidential)

2. Holleman, E. C., Armstrong, N. A., and Andrews, W. H., "Utilization of the Pilot in the Launch and Injection of a Multi-Stage Orbital Vehicle," Paper presented at the IAS 28th Annual Meeting, New York, January 26, 1960.

3. Muckler, F. A., Hookway, R. O., and Burke, H. H., "Manned Control of Large Space Boosters," Trans. Seventh Symposium on Ballistic Missile and Space Technology, USAF, Vol. II, August 15, 1962, pp. 149-188.

4. Muckler, F. A., and Obermayer, R. W., "The Use of Man in Booster Guidance and Control," Paper presented at the AIAA Guidance and Control Conference, August 1963.

5. Marshall Space Flight Center, "Saturn V Launch Vehicle Design Data Book," Marshall Document IN-P&VE-V-62-6, 1962.

F. ADAPTIVE PHASE-STABILIZATION OF BENDING FOR LARGE AEROSPACE VEHICLE BOOSTERS by Michael T. Borelli and Hans H. Hosenthien

This article describes three adaptive control techniques being evaluated to provide phase-stabilization for the low frequency structural bending modes of large aerospace vehicle boosters, such as Saturn V. A detailed discussion of how the phase characteristics of an adaptive tracking notch filter can be utilized to obtain phase-stabilization is presented. Two other adaptive control schemes, the model reference system and the phase-sensing system, are discussed. Results from an analog computer simulation of the three adaptive systems are reviewed and briefly evaluated. All systems performed well under wide variations in the first bending mode parameters.

INTRODUCTION

Stabilization of closed-loop feedback control systems to undesirable structural bending mode signals as found in large aerospace vehicle boosters can be accomplished by either gain- or phase-stabilization techniques. Gain-stabilization prevails when the energy fed back to the structure through the control loop is less than that required to sustain the structure in oscillation. Phase-stabilization, obtained by proper phasing of the feedback signal, effectively removes energy from the structure and therefore forces the system back to equilibrium faster than gain-stabilization. An example of a gain-stabilized system is the adaptive tracking notch (ATN) filter developed by the Astrionics Laboratory for Saturn V (Refs. 1 and 2). Although the performance of this system was highly satisfactory, it was decided that an adaptive phase-stabilized system would be more desirable because of the resulting reduced structural loading of the vehicle. The gain and phase characteristics of the ATN filter are very well suited for use in a phase-stabilized system, but certain modifications are necessary. Two other promising phase adaptive control schemes, the model reference system and the phase-sensing system, were also investigated.

The authors would like to acknowledge the contributions of the personnel of the Martin-Marietta Company of Denver, Colorado, who proposed the model reference and phase-sensing approach and conducted the investigations. Most of the data used in this report were taken from their study (Ref. 3).

NOMENCLATURE

C	capacitor
K	integrator gain, constant
K_A	accelerometer gain
K_D	displacement gain
K_R	rate gain
K_{AD}	adaptive gain
K', K''	constants
M	generalized mass
q	dynamic pressure
R	resistor
R'	thrust of control engines
S	complex variable
X_β	station of engine gimbal
$Y_{(x)}$	normalized deflection at station x resulting from the first bending mode
$Y'_{(x)}$	normalized slope at station x resulting from the first bending mode
$\ddot{Z}$	acceleration normal to reference
Z_{S2}	slosh mass displacement of second mode
α	angle of attack
α_ω	angle of attack caused by wind
β	engine deflection
ζ_{Bi}	modal damping of the ith mode
η_i	generalized displacement of the ith mode
Θ	phase angle
τ	time constant

$\ddot{\tau}_{(x)}$	accelerometer signal at station x
φ_c	attitude command angle
$\varphi_{(x)}$	attitude signal at station x
$\dot{\varphi}_{(x)}$	attitude rate signal at station x
φ_{Bi}	attitude angle resulting from ith bending mode
ω_c	control frequency
ω_o	operating frequency
ω_{Bi}	modal frequency of the ith mode
$\dot{\delta}, B'_c$	outputs from the adaptive band-pass filters
ϵ	error signal

ADAPTIVE TRACKING OFFSET NOTCH FILTER

The adaptive tracking offset notch (ATON) filter utilizes the mechanization of the ATN filter with some additional circuits to obtain the offset notch. A complete description of the ATN filter operation is given in References 1 and 2, but for convenience a brief description will be given here. The basic operation of the ATN filter is to construct a signal with the same frequency as the bending mode signal to be suppressed, but with a 180-degree phase relationship. The amplitude of the fundamental component of this constructed signal is equal to the amplitude of the bending mode signal. When these two signals are summed, as shown in Figure 1, the bending mode signal is cancelled.

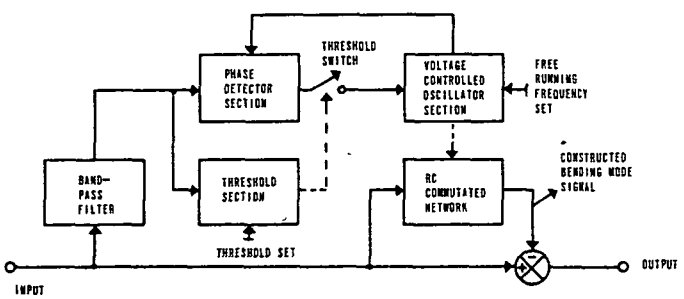


FIGURE 1. ADAPTIVE TRACKING NOTCH FILTER

The input to the ATN filter is a composite of rigid body signals, sloshing signals, bending mode signals, etc. The band-pass filter is tuned for the range of the bending mode signal to be suppressed. When the output signal from the band-pass filter exceeds a set threshold value, a switch is closed connecting the phase detector circuit to the voltage controlled oscillator (VCO). With the switch closed, phase-locked operation is achieved and the VCO is thereby forced to oscillate at the frequency of the output signal of the band-pass filter. This provides tracking of the bending mode frequency throughout the first stage flight. When the switch shown in Figure 1 is open, the VCO is set to oscillate at a frequency approximately in the middle of the expected bending mode frequency range; this is called the free running frequency.

The output of the VCO section is a four-phase signal at the bending mode frequency. This four-phase signal, with a 0, 45, 90, and 135-degree phase relationship, is used to commutate four capacitors in the RC commutated network. The RC commutated network for this application is an operational amplifier with four commutated capacitors connected in series in a feedback loop. A second resistor feedback loop is also connected around the amplifier.

The operation of the RC commutated network entails demodulation of the input bending mode signal, smoothing and suppression of the other frequencies present, and then modulation of a signal at the commutated or bending mode frequency. The output of the RC commutated network has a staircase waveform with the amplitude of the fundamental equal to the amplitude of the bending mode signal. The higher harmonics present in the staircase wave (seventh and higher) are suppressed with a conventional filter designed to handle the higher bending mode signals.

A typical frequency response for the ATN filter is shown in Figure 2. This response is obtained by holding the frequency of the input signal to the band-pass filter constant, and varying the frequency of the input signal to the RC commutated network. Figure 2 shows that the first bending mode frequency (ω_{B_1}) is suppressed by a factor of at least 100 to 1; this attenuation is obtained throughout its range. With the notch at the first bending mode frequency, little effect can be noted on the phase and amplitude characteristics of the control frequency (ω_c) and the second bending mode frequency (ω_{B_2}).

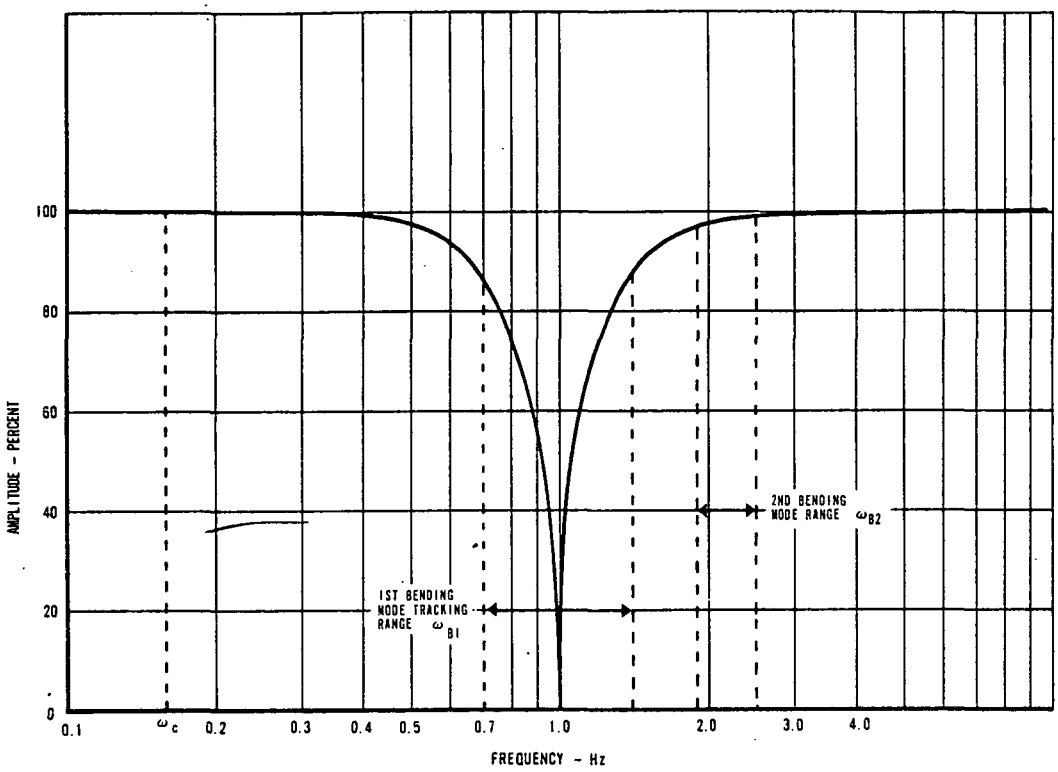


FIGURE 2A. ATN FILTER AMPLITUDE FREQUENCY RESPONSE

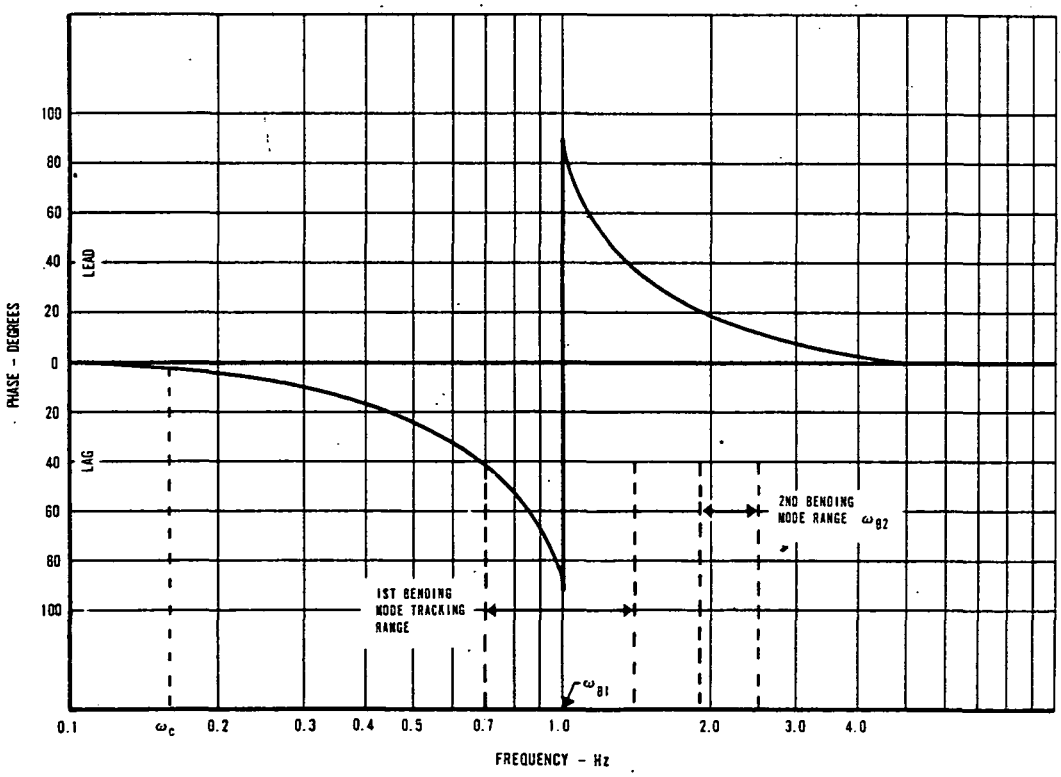


FIGURE 2B. ATN FILTER PHASE FREQUENCY RESPONSE

Since the shape of the phase frequency response curve is essentially the same throughout the range of ω_{B1}, offsetting the notch from the actual bending frequency yields a fixed amount of phase shift at ω_{B1}, resulting in an adaptive tracking offset notch filter. The notch offset at ω_{B1} has little effect on the gain or phase at the control frequency or at higher bending mode frequencies.

Studies conducted by Hart, et al (Ref. 3) established that -10 to -15 percent offset for the ATON filter gave a satisfactory amount of phase lead to phase-stabilize the first bending mode of Saturn V. An offset in this range will increase the effective structural damping from an initial value of 0.5 percent to a value of 4 to 10 percent, depending on the system parameters. A control system utilizing the ATON filter can tolerate a five-times-nominal increase in slope of the first bending mode shape of the Saturn V vehicle at the selected station of the control rate gyro. This is comparable to the gain-stabilized system that utilized the ATN filter (Refs. 1 and 2).

Two mechanizations for providing the frequency offset for the ATON filter have been investigated; one utilizes a second VCO (Fig. 3), the other utilizes digital techniques to obtain an incremental frequency offset. In Figure 3, VCO-I is utilized in a phase-locked loop to track the bending mode frequency. VCO-II has an input that shifts the frequency to the desired offset value. One objection to this system is the essentially open loop operation of VCO-II; however, open loop operation can be eliminated by using one VCO and additional digital circuits to obtain an incremental frequency offset. A 12.5 percent frequency offset can be obtained with a reasonable number of additional digital circuits (diodes and flip flops). By adjusting the time constant (R and C parameters) of the RC commutated network, a rather wide variation in phase response curves can be obtained (Fig. 4) at the incremental frequency offset.

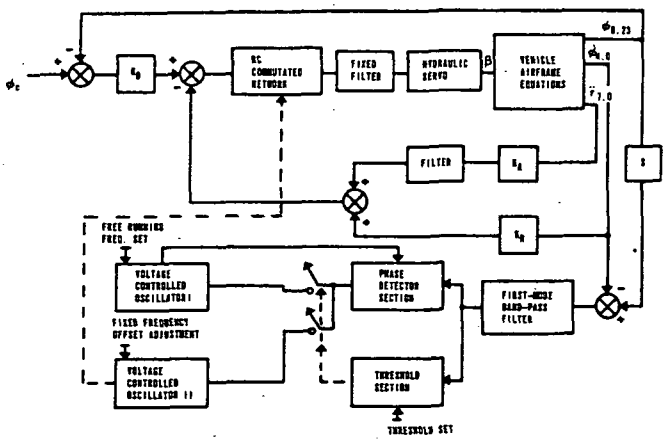


FIGURE 3. OFFSET NOTCH FILTER ADAPTIVE CONTROL SYSTEM

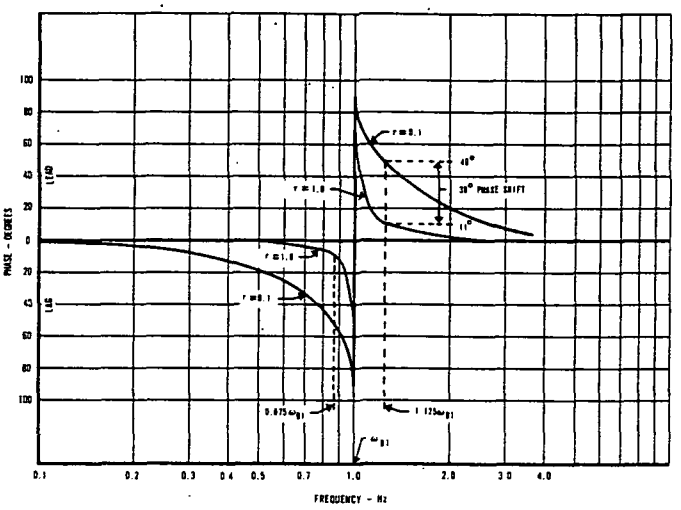


FIGURE 4. PHASE SHIFT ATTAINABLE WITH VARIATIONS IN τ OF COMMUTATED NETWORK

MODEL REFERENCE AND PHASE-SENSING ADAPTIVE SYSTEMS

The model reference adaptive control system of Figure 5 uses a zero model (Ref. 4). Any first bending mode signal from the vehicle is an error signal; therefore, the signal from the model, and also the model itself, is zero. The adaptive gain from Figure 5 is

$$K_{AD} = K \int \frac{\partial \epsilon}{\partial K_{AD}} \epsilon \, dt \qquad (1)$$

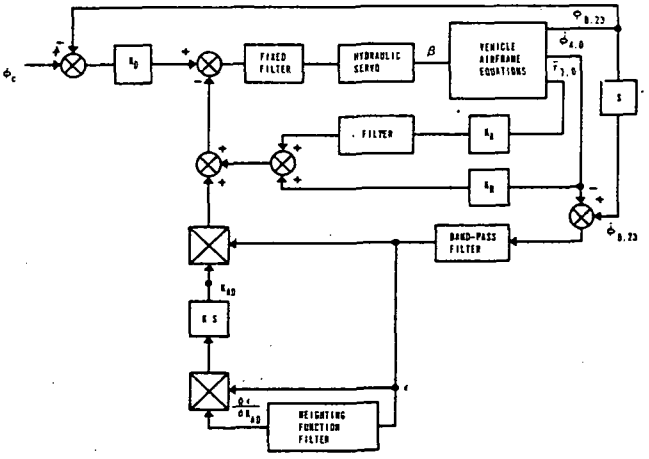


FIGURE 5. MODEL REFERENCE ADAPTIVE CONTROL SYSTEM

where ϵ is the error or the first bending mode signal. The error signal is obtained by subtracting two attitude rate signals to eliminate the rigid body signals. A band-pass filter, tuned to the first bending mode frequency range, suppresses the higher bending mode signals.

The expression for K_{AD} can be shown to be the integral error squared. Minimizing this integral is quite often used as the performance index for adaptive control systems. In this application however, it is used to force the system to oscillate at some closed-loop operating frequency ω_o. The value of ω_o is selected by making a parametric study of the first bending mode as a function of K_{AD} for different flight times. Figure 6 shows the root locus resulting from varying K_{AD} for three values of the first bending mode frequency; nominal (ω_{B_1}), 1.2 times nominal ($1.2\,\omega_{B_1}$), and 0.8 times nominal ($0.8\,\omega_{B_1}$). The root loci for variations in the slope of the first bending mode shape are shown in Figure 7. The selected operating frequency gave maximum effective closed loop damping for nominal conditions.

NOTE: 1. FLIGHT TIME=80 SECONDS
2. ADAPTIVE LOOP BAND-PASS FILTER= $\dfrac{s^2}{10\left[\left(\frac{s}{7.0}\right)^2+\frac{2(1.0)s}{7.0}+1\right]\left[\left(\frac{s}{5.0}\right)^2+\frac{2(0.7)s}{5.0}+1\right]}$

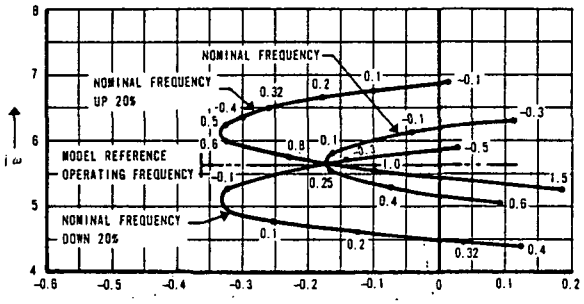


FIGURE 6. FIRST BENDING MODE ROOT LOCUS (K_{AD} VARIABLE) FOR MODEL REFERENCE SYSTEM WITH FIRST MODE FREQUENCY VARIED

NOTE: 1. FLIGHT TIME=80 SECONDS
2. ADAPTIVE LOOP BAND-PASS FILTER= $\dfrac{s^2}{10\left[\left(\frac{s}{7.0}\right)^2+\frac{2(1.0)s}{7.0}+1\right]\left[\left(\frac{s}{5.0}\right)^2+\frac{2(0.7)s}{5.0}+1\right]}$

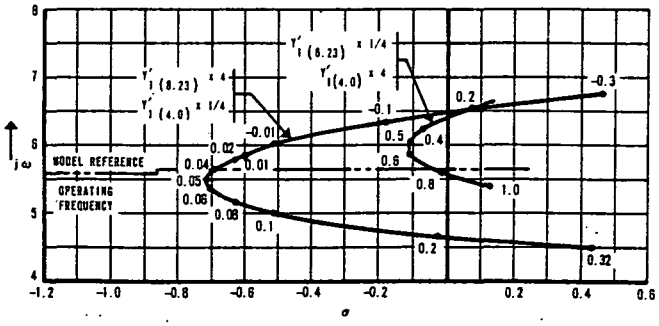


FIGURE 7. FIRST BENDING MODE ROOT LOCUS (K_{AD} VARIABLE) FOR MODEL REFERENCE SYSTEM WITH FIRST MODE SLOPE VARIED

When the operating frequency has been determined, the weighting function filter is then designed with a phase shift of 90 degrees at the operating frequency. If the error signal ϵ is assumed as

$$\epsilon = K \sin \omega_{B_1} t \qquad (2)$$

then the output from the weighting function filter is

$$\frac{\partial \epsilon}{\partial K_{AD}} = K' \sin (\omega_{B_1} t + \Theta) \qquad (3)$$

where Θ is the phase shift at ω_{B_1}. Their product is

$$\frac{\partial \epsilon}{\partial K_{AD}} \cdot \epsilon = K'' [\cos \Theta - \cos(2\omega_{B_1} t + \Theta)]. \qquad (4)$$

The first term ($\cos \Theta$) is a dc signal and, when integrated, results in the adaptive gain K_{AD}. The adaptive gain will change until Θ equals 90 degrees, or $\omega_{B_1} = \omega_o$; the first bending mode poles are at the operating frequency. The second term of equation 4 is attenuated in the integrator and its effect is not considered significant.

The phase-sensing adaptive control system is similar to the model reference system just discussed because it utilizes an adaptive gain K_{AD} to control the location of the closed-loop first mode bending poles. However, the mechanization of the phase-sensing system (Fig. 8) is considerably different from the model

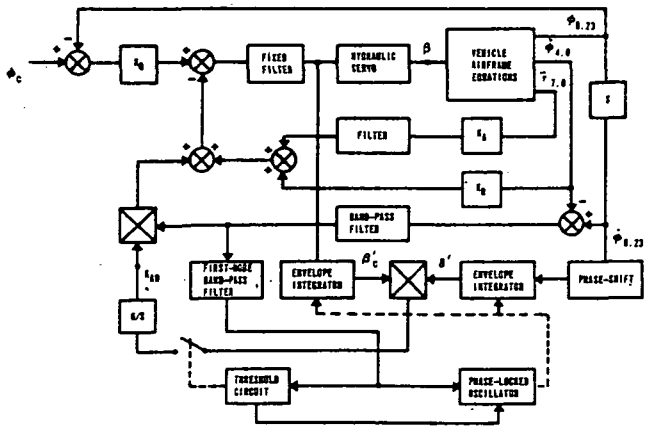


FIGURE 8. PHASE-SENSING ADAPTIVE CONTROL SYSTEM

reference system. The envelope integrators shown in Figure 8 are the same as the RC commutated networks of the ATN filter (Fig. 1) and, when used as shown, exhibit a highly selective band-pass characteristic. Also, unlike the model reference system, the phase-sensing system does not force the first mode bending poles to some preselected operating frequency, but rather enforces a 90-degree phase relationship between the engine deflection angle β and the attitude angle caused by bending, φ_{B1}. A 90-degree phase relationship between β and φ_{B1} will increase the effective damping of the system and thus remove energy from the oscillating system.

For analytical purposes, the structural bending modes are considered as simple, linear, lightly damped spring-mass systems. An equation representing the first bending mode can be written as

$$\ddot{\varphi}_{B1} + 2\zeta_{B1}\omega_{B1}\dot{\varphi}_{B1} + \omega^2_{B1}\varphi_{B1} = -\frac{R'Y_{(x\beta)}Y'_{(RG)}}{M}\beta \qquad (5)$$

where ζ_{B1} = modal damping

ω_{B1} = modal frequency

R' = thrust of control engines

$Y_{(x\beta)}$ = deflection at the gimbal

$Y'_{(RG)}$ = slope at the rate gyro

M = generalized mass.

The response of such a second-order system (with small ζ_{B1}) to an impulse will be a slightly damped sinusoidal oscillation. By rewriting equation 4 as

$$\ddot{\varphi}_{B1} + \omega^2_{B1}\varphi_{B1} = -2\zeta_{B1}\omega_{B1}\dot{\varphi}_{B1} - \frac{R'Y_{(x\beta)}Y'_{(RG)}}{M}\beta \qquad (6)$$

and considering the product of $Y_{(x\beta)}$ and $Y'_{(RG)}$ positive, it is seen that the effective damping force is increased when β is in phase with $\dot{\varphi}_{B1}$. For the assumed sinusoidal oscillations, a 90-degree phase relation must exist between φ_{B1} and β in order that $\dot{\varphi}_{B1}$ be in phase with β.

A first bending mode parametric study, with the gain K_{AD} variable, yields root loci of similar shape as the model reference system. The closed loop frequency of the first mode bending poles corresponds closely to the open loop frequency as can be seen in Figures 9 and 10.

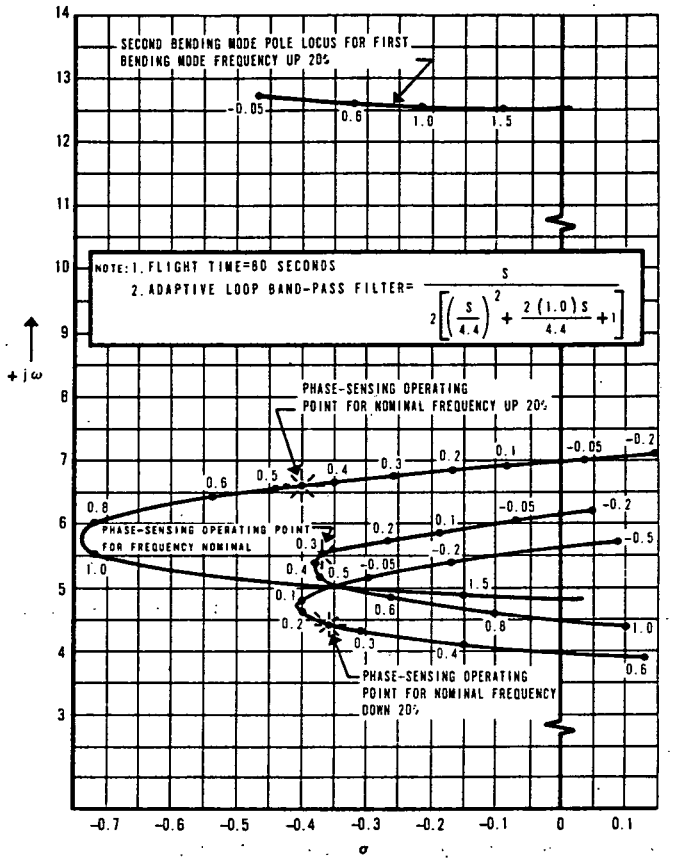

SECOND BENDING MODE POLE LOCUS FOR FIRST BENDING MODE FREQUENCY UP 20%
NOTE: 1. FLIGHT TIME=80 SECONDS
2. ADAPTIVE LOOP BAND-PASS FILTER= $\frac{S}{2\left[\left(\frac{S}{4.4}\right)^2 + \frac{2(1.0)S}{4.4} + 1\right]}$
PHASE-SENSING OPERATING POINT FOR NOMINAL FREQUENCY UP 20%
PHASE-SENSING OPERATING POINT FOR FREQUENCY NOMINAL
PHASE-SENSING OPERATING POINT FOR NOMINAL FREQUENCY DOWN 20%
+jω
σ

FIGURE 9. FIRST BENDING MODE ROOT LOCUS (K_{AD} VARIABLE) FOR PHASE-SENSING SYSTEM WITH FIRST MODE FREQUENCY VARIED

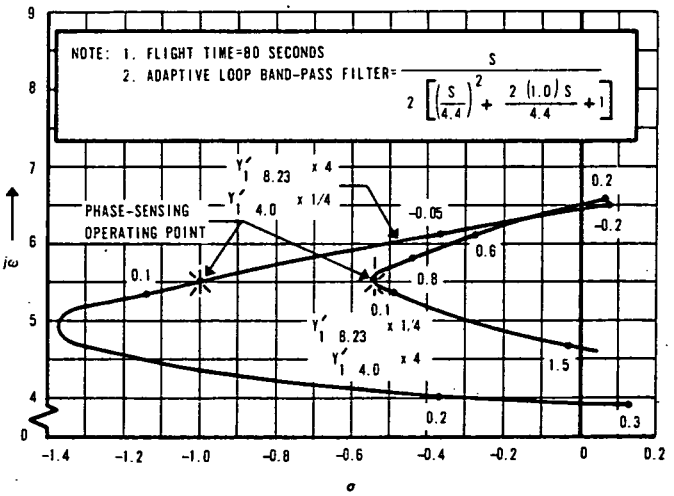

NOTE: 1. FLIGHT TIME=80 SECONDS
2. ADAPTIVE LOOP BAND-PASS FILTER= $\frac{S}{2\left[\left(\frac{S}{4.4}\right)^2 + \frac{2(1.0)S}{4.4} + 1\right]}$
PHASE-SENSING OPERATING POINT
Y'1 8.23 x 4
Y'1 4.0 x 1/4
Y'1 8.23 x 1/4
Y'1 4.0 x 4
jω
σ

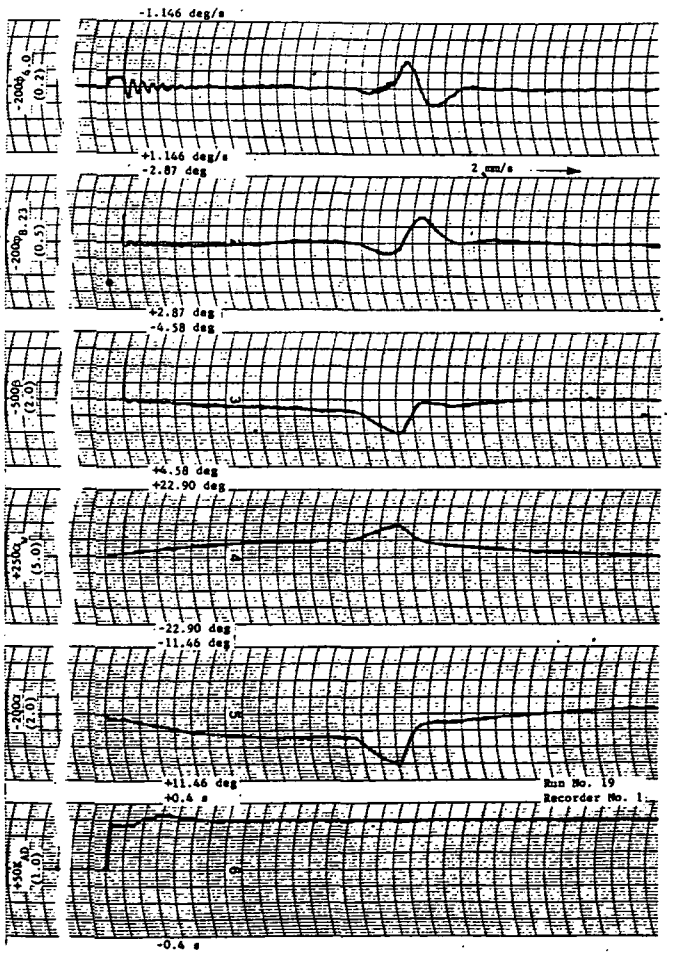
FIGURE 10. FIRST BENDING MODE ROOT LOCUS (K_{AD} VARIABLE) FOR PHASE-SENSING SYSTEM WITH FIRST MODE SLOPE VARIED

ANALOG SIMULATION

An analog computer simulation was conducted by the Martin Company (Ref. 3) to compare the performance of the three phase-adaptive control systems using the pitch axis of a simplified version of the Saturn V vehicle.

The simulation included the first two body bending modes and one propellant slosh mode. Time varying coefficients were used and the time of flight studied was the critical period from 40 to 120 seconds, or the region of maximum dynamic pressure. Specifications for the Saturn V provide for an αq product of less than 35,800 kg-deg/m^2 which yields a maximum angle-of-attack (α), in the region of maximum dynamic pressure, of 9.8 degrees. Maximum engine deflection (β) was specified as ±5.0 degrees. The three systems were studied for:

1. First bending mode slope changes of $4 \times Y_1'$ (8.23), $\frac{1}{4} \times Y_1'$ (4.0) and $\frac{1}{4} \times Y_1'$ (8.23), $4 \times Y_1'$ (4.0).



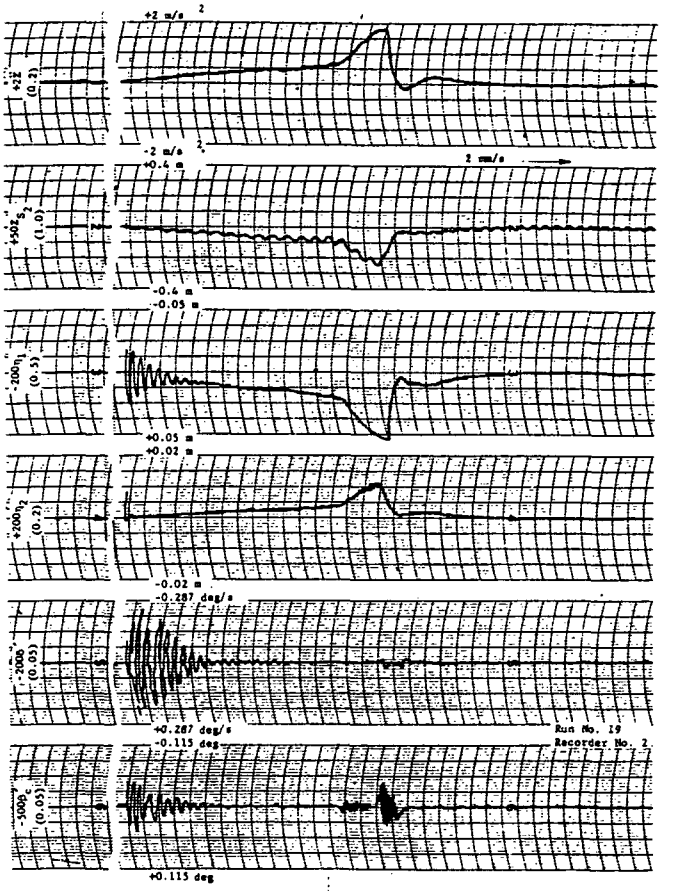


FIGURE 11. PHASE-SENSING SYSTEM, TIME OF FLIGHT 40-120 SECONDS

2. First bending mode frequency variations of ± 20 percent.

3. Adaptive loop isolation from other modes (second bending mode and propellant slosh mode).

4. Combinations of items 1 and 2 above.

Twenty-nine analog computer runs were made of the three systems. The only run resulting in unstable operation was the model reference system for conditions of $\frac{1}{4} \times Y_1'$ (8.23), $4 \times Y_1'$ (4.0), and $\omega_{B1}/1.2$. For all other parameter variations and initial conditions, the three systems were stable. Initial conditions included slosh wave heights up to 0.43 meter and first and second bending mode deflections up to 0.037 meter. These runs showed that the presence of a slosh mode and second bending mode did not upset the adaptive operation of these systems.

A typical analog recording of the phase-sensing system is shown in Figure 11. For this run, the first

bending mode frequency was decreased ($\omega_{B1}/1.2$), and an initial condition in the deflection of the first bending mode ($\eta_1 = 0.022$ m) was imposed.

CONCLUSIONS

The three phase-adaptive control systems provide at least six times greater damping to the first bending mode signal than is present because of the vehicle's inherent structural damping. All systems performed well under wide variations in the first bending mode parameters. Analog simulations that included one sloshing mode and the second bending mode did not indicate any adverse operation of any of the systems. Mechanization of the three systems does not appear to present any great difficulties.

Although none of the three systems exhibited a clear superiority over the others, the ATON filter and phase-sensing systems consistently provided greater phase margins than the model reference system. Also, the model reference system did not handle the ±20 percent variation in the first bending mode frequency as well as the other two systems. A detailed investigation of this deficiency will be made as it could present a serious limitation to the system. This is particularly true if the variation in frequency is combined with a variation in the bending mode slope.

Additional studies of the three systems discussed here are planned. An effort to select the most promising system based on performance and mechanization simplicity will be made, and a breadboard of the best system will be built.

REFERENCES

1. Hosenthien, Hans H., and Borelli, Michael T., "An Adaptive Tracking-Notch Filter for Suppression of Structural Bending Signals of Large Space Vehicles," Astrionics Research and Development Report No. 1, NASA TM X - 53,000, October 1, 1963.

2. Hart, W. G., and McGowan, G. F., "Advanced Control System for Saturn V Configuration, Phase I Final Report," NASA - CR - 63 - 24, Martin Company, Denver, Colorado, August 1963 (Confidential).

3. Hart, W. G., McGowan, G. F., and Vandenberg, F. A., "Advanced Control System for Saturn V Configuration, Phase II Final Report," NASA - CR - 63 - 108, Martin Company, Denver, Colorado, November 1963 (Confidential).

4. Kezer, A., Hofmann, L. G., and Engel, A. G., Jr., "Application of Model-Reference, Adaptive Control Techniques to Provide Improved Bending Response of Large Flexible Missiles," Sixth Symposium on Ballistic Missile and Aerospace Technology, Massachusetts Institute of Technology, Cambridge, Massachusetts, June 1961.

G. A SYSTEM FOR ACQUISITION AND REDUCTION OF DYNAMIC TEST DATA OF LARGE AEROSPACE VEHICLES

by Frank Hay

Dynamic test data in analog form are digitized by a successive approximation technique and placed in format for direct computer entry. Phase distortion normally introduced by such acquisition is minimized by employing simultaneous sample-and-hold. The data are reduced by a computer program based on the Fischer-Hinnen method of analyzing a periodic wave by fitting a finite trigonometric sum to a set of recorded digitized values.

This technique permits the acquisition and reduction of an abundance of test data at rates commensurate with the test schedule. The results are used to determine the vehicle dynamic response characteristics required for control systems design.

INTRODUCTION

Accurate determination by dynamic testing of the structural gain and phase characteristics of large aerospace vehicles, such as the Saturn, is required in the design of the control system. Figure 1 shows a typical vehicle configuration and locations of the sensors for determination of structural characteristics with respect to a given orientation. These sensors produce analog output voltages that are properly scaled for the recording devices through instrumentation hardware such as amplifiers, demodulators, etc. Voltage outputs of the load cell that measures the shaker force, the oscillator that drives the shaker, and the sensors are recorded in analog form (Fig. 2) on strip recorders. These records are used to visually monitor the test in the event of operation errors, equipment malfunctions, local disturbances, or structural nonlinearities.

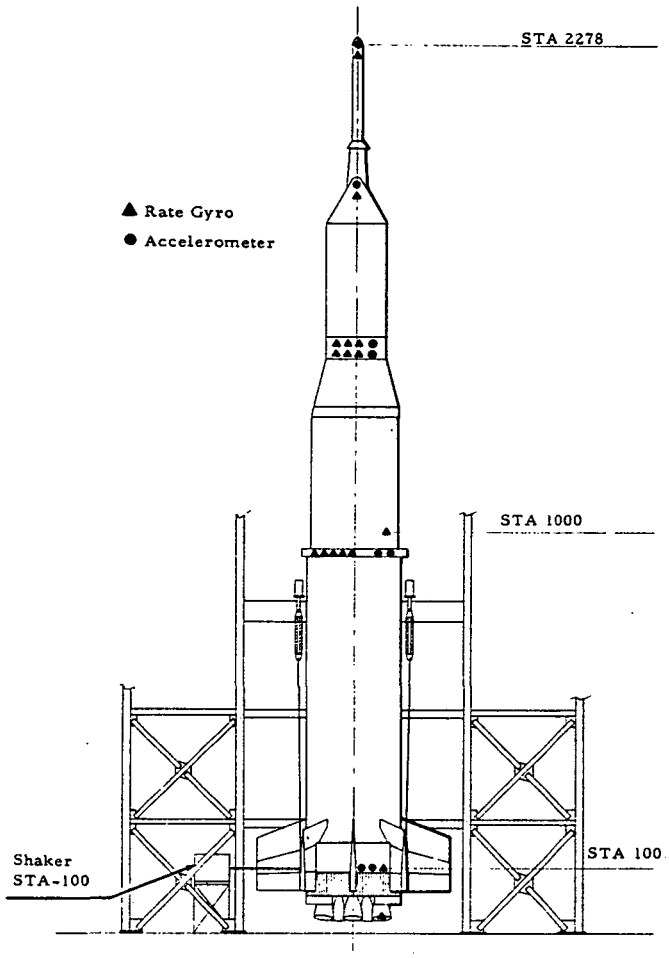


FIGURE 1. DYNAMIC TEST CONFIGURATION, SA-6, SATURN I

From this data, transfer functions (gain and phase characteristics) between any two of a maximum of thirty sensors on the vehicle, or between any of these sensors and the shaker force sensed by a load cell, must be computed. A typical gain versus frequency plot or transfer function is shown in Figure 3.

This article describes a system for acquiring and reducing the large quantities of data generated by dynamic tests. Digital techniques are applied to accomplish this task at rates commensurate with the test schedule.

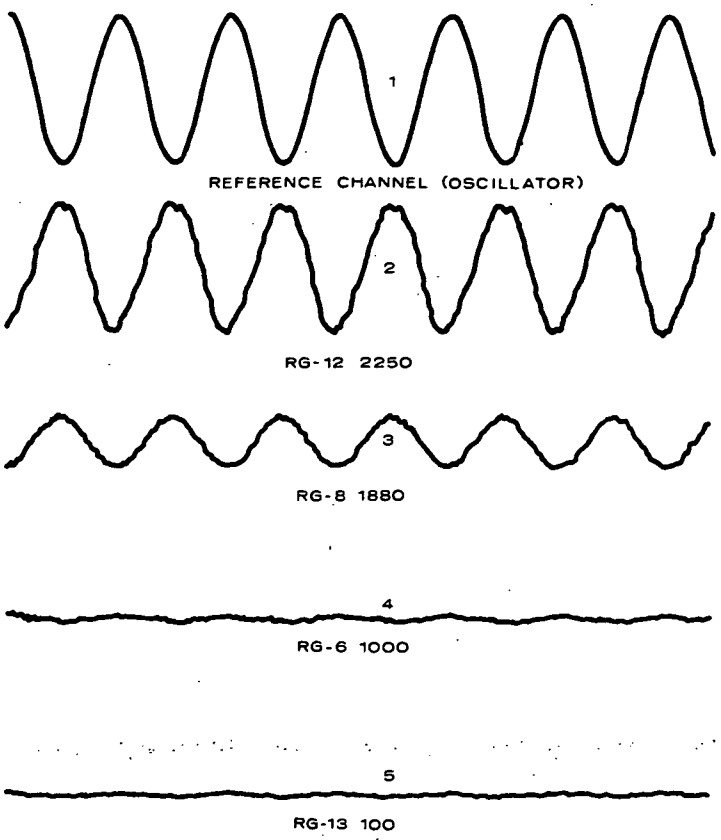


FIGURE 2. TYPICAL OUTPUTS OF ANALOG SENSORS (RATE GYROS)

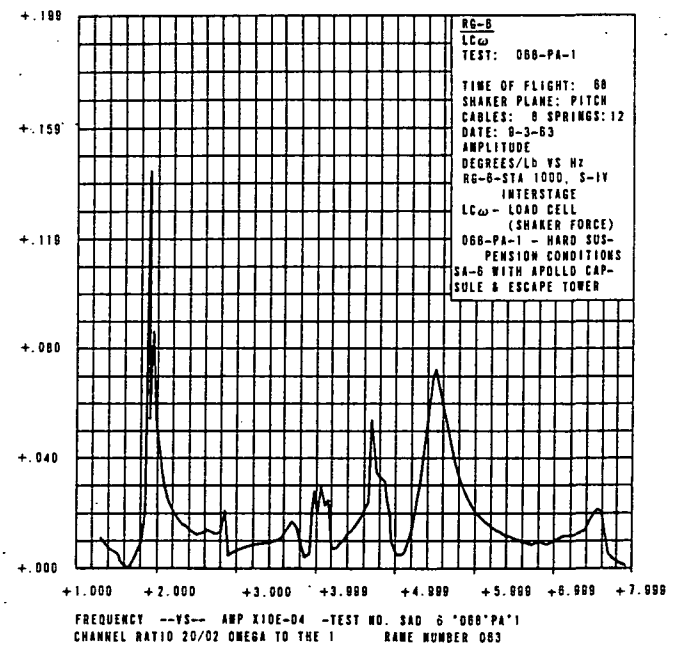


FIGURE 3. PLOT OF TRANSFER FUNCTION BETWEEN LOAD CELL AND ATTITUDE CONTROL SENSOR

NOMENCLATURE

V_A	amplitude of sampled analog input voltage
V_R	ladder reference voltage
T_{RD}	character transfer rate to the magnetic tape
v	speed at which tape moves through write head
D	tape character density
W_R	analog to digital converter word rate
t_w	analog to digital converter word time
t_h	sample-and-hold time
t_c	sample-and-hold charging time
t_s	time between samples of analog voltage
SR	sampling rate
A_C	charging amplifier for sample-and-hold capacitors
A_H	holding amplifier for sample-and-hold capacitors
V_o	sample-and-hold output voltage
C_g	number of characters to be stored during gap time
l_g	length of tape gap
x_i	sampling increment
y_i	ordinate at the sampling increment
d_n	digital number corresponding to instantaneous analog input
R_I	input impedance of hold amplifier
C_h	information channel
m	number of equal division of the period T

N number of digitized values between the zero crossover points of successive positive slopes of the oscillator channel

p_i interpolation constant

α phase angle resolution

w_t tape characters per computer word

SYSTEM CONCEPT

The control system designer is primarily interested in the fundamental frequency response of the vehicle. However, the analog sensor voltages will contain the fundamental or driving force frequency plus harmonics resulting from distortion caused in part by random errors inherent in acquisition and recording equipment, high frequency electrical transients, sensor noise, and nonlinearities of the vehicle structure (Fig. 2). Therefore, for the steady-state condition (no structural transients), the vehicle fundamental frequency response to a driving force,

$$F = F_1 \sin(\omega t + \theta) \quad (1)$$

will be of the form

$$R = R_1 \sin(\omega t + \varphi) \quad (1a)$$

(θ and φ are referenced to some arbitrary origin) after performing a Fourier analysis on the analog voltages. Higher harmonics are easily extracted if desired. A reference or origin for performing the Fourier analysis on the load-cell and sensor waveforms is provided by recording the shaker-driving oscillator voltage on information channel 1. This voltage is chosen as a reference because its output waveform is frequency stable and practically free of noise and harmonic distortion.

The data, reduced to this form, are presented to the control systems designer for transfer function selection. It is then routine to compute the transfer function between any two sensors simply by dividing the amplitudes and subtracting the phase angles.

A block diagram of a system to accomplish this is shown in Figure 4.

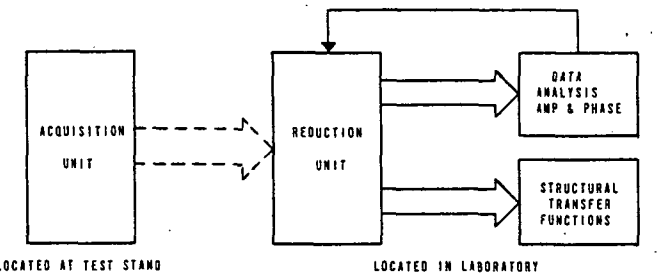


FIGURE 4. SYSTEM BLOCK DIAGRAM

A high-speed digital data acquisition unit is used to acquire, digitize, place in format, and record the analog outputs of the sensors at a real time rate. This unit is located at the dynamic test stand and records the acquired data on magnetic tape for later entry into the reduction unit computer in the laboratory (Fig. 4). The data consist of test identification and digitized analog sensor voltages. Amplitude accuracy is assured by using the number of weighted approximations dictated by the desired control systems accuracy in a successive approximation technique of analog to digital conversion. Precise cross-channel phase correlation is maintained through the use of simultaneous sampling and holding of all channels. The time interval between samples is accurately controlled, allowing responses to very small frequency changes to be recorded. This feature is advantageous when determining vehicle resonances as illustrated by Figure 3. This graph shows a sharp resonance whose entire peak occurred in a 0.15 Hz increment between 1.85 and 2.00 Hz.

A general purpose digital computer with appropriate accessories is operated under optimized program control to reduce these data. A part of the program, based on the Fischer-Hinnen method of determining the Fourier coefficients of a periodic function, calculates and prints out the desired result.

The main advantages of this system are speed and accuracy. The system is designed for the overall accuracy required by the control systems designers. Producing complete results from acquired data of an average four hour test within three to six hours (depending on frequency range covered during the test) is evidence of its speed. This compares to several days to reduce similar quantities of data manually. Other advantages include flexibility in reduction program, repeatability of results, and ease of updating with regard to the number of channels and amplitude accuracy for testing vehicles having more complex structures.

SYSTEM DESCRIPTION

Data Acquisition Unit

A block diagram of the data acquisition unit is shown in Figure 5. The acquisition unit was designed and built for acquiring data in the frequency range of 0 to 30 Hz. It has a capacity of 32 information channels and simultaneously holds the instantaneous level of data of each channel for subsequent digitizing, placing in format, and recording on magnetic tape for direct compatible entry into the reduction unit computer. The acquisition unit samples the inputs at a rate of 612 samples/s/channel. The analog voltages of ± 4.096 volts maximum are converted to digital numbers that represent the instantaneous level of data with ± 0.15 percent accuracy and maintain a cross-channel phase correlation within ± 1.0 degree at 30 Hz.

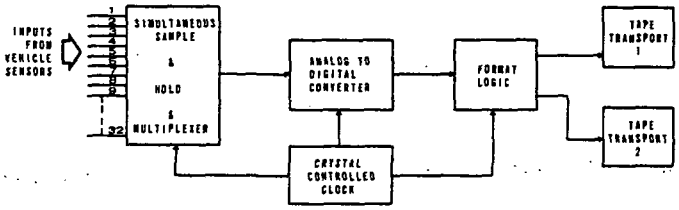


FIGURE 5. DATA ACQUISITION UNIT

This unit will be considered only from the system approach. The requirements dictated by system consideration will be investigated for each sub-unit shown in block form in Figure 5. The limitations placed upon system performance by these units will also be considered. The investigations dictate the specifications from which the circuit design, logic design, and fabrication were accomplished.

The sub-units shown in Figure 5 are included for the purposes of:

a. Converting analog signals obtained from vehicle sensors to equivalent digital numbers for later utilization of high-speed digital computers.

b. Providing amplitude accuracy compatible with the desired system accuracy (Ref. 1).

c. Placing the digital information in format for direct entry into a computer, thereby eliminating data handling errors inherent in intermediate processing (Ref. 2).

d. Minimizing phase distortion introduced during acquisition.

e. Achieving a high sampling rate so data can be acted upon directly without elaborate curve-fitting programs to effect recovery of the digitized waveform.

f. Providing accurate sample intervals for use as a computer time reference.

The analog voltages (Fig. 2) are converted to equivalent digital numbers by an analog to digital converter using a successive approximation technique (Ref. 3). The resultant coded voltage is a function of the ladder mechanization shown in Figure 6.

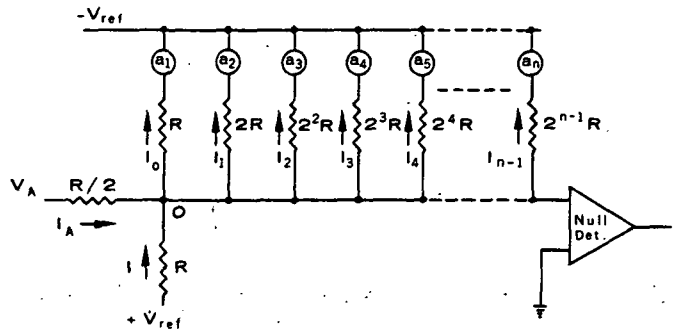


FIGURE 6. ANALOG TO DIGITAL CONVERTER LADDER

V_A is considered to be constant since it represents the instantaneous data voltage and is held by the sample-and-hold unit until all channels are converted to a corresponding digital number. (The effect on amplitude accuracy of the sample-and-hold will be discussed later.)

Summing the currents about node 0 (neglecting ladder loading) and solving for V_A yield

$$V_A = V_R \left(\frac{dn - 2^{n-1}}{2^n} \right) \qquad (2)$$

where $dn = a_1 2^{n-1} + a_2 2^{n-2} + a_3 2^{n-3} + a_4 2^{n-4} + \ldots a_n 2^0$; and $a_1, a_2 \ldots a_n$ can have values of 1 or 0 only. Tolerances on ladder supply voltages and resistors were not considered in expression 2 since compensation for correction of these errors is built in. The system analog to digital (A/D) converter has n = 11 bits (resulting in a resolution accuracy of 1 part in 2047 or .049 percent) and a linearity of .0125 percent (Ref. 5). Expression 2 is used to calculate the instantaneous sensor amplitudes in the reduction program.

The rate at which the analog data can be sampled is limited by the output tape transports. Two tape units are provided for the purpose of continuous data acquisition. When the tape on unit 1 has been completely written, the data are automatically transferred to tape unit 2.

The character transfer rate for a magnetic tape unit is the product of the tape speed, v, and the tape character packing density, D. It is given by

$$T_{RD} = vD. \tag{3}$$

The tape transports in this system are standard 7 track, high-transfer-rate units with an operating speed of 286 cm/s at a packing density of 218.7 characters/cm. This gives a transfer rate of $T_{RD} = (286)(218.7) = 62,500$ characters/s. Thus, one 7-bit character is written on tape every 16 μs.

The reduction computer has a 20-bit word length. Therefore, only one A/D converter word can be stored per computer word for an A/D converter with 11-bit accuracy. This computer word is written in a special tape format of three characters; therefore, the acquisition unit must have a word transfer rate, W_R, of

$$W_R = \frac{vD}{w_t}. \tag{4}$$

For this system, $W_R = \frac{62,500}{3} \frac{\text{wds}}{\text{s}}$.

The acquisition word time is the reciprocal of the word rate

$$t_w = \frac{1}{W_R} = \frac{3}{62,500} = 48\,\mu s \tag{5}$$

Each word on tape represents one digitized point for only one function.

For 32 functions, the time t_h is

$$t_h = 32\, t_w = 1536\ \mu s \tag{6}$$

where t_h is the time required for the sample-and-hold unit to hold. The sample-and-hold unit used in this system has a hold-time-to-charge-time ratio of 20 to 1. Therefore, the charge time is

$$t_c = \frac{1}{20}\ 1536 = 76.8\ \mu s.$$

This is significant since a return to the hold condition must occur at least 76.8 μs after sampling a new voltage. Therefore, two word times (96 μs) are assigned for the purpose of allowing the sample-and-hold capacitors to charge. This gives a sampling period of

$$t_s = t_h + 2t_w = 1632\ \mu s \tag{7}$$

which is generated very accurately from the crystal-controlled clock located in the A/D converter and is used in the reduction program for frequency calculation. The sampling rate, SR, is the reciprocal of this time

$$SR = \frac{1}{t_s} = 612 \text{ samples/s/channel}. \tag{8}$$

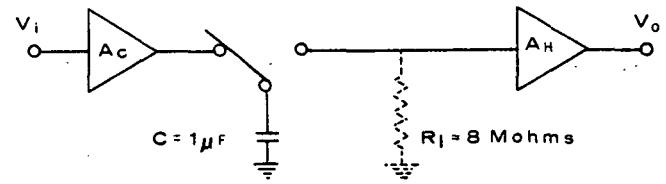


FIGURE 7. SIMPLIFIED SAMPLE-AND-HOLD UNIT

The effect of the sample-and-hold unit (Fig. 7) on the amplitude is investigated in regard to system accuracy. The capacitor C (1 μF) of a typical channel charges to V_{o1} (sample time) through the amplifier A_C and is switched at the hold command time to the input of the hold amplifier A_H. C is then shunted during hold time t_h by the input impedance, R_I, of A_H.

R_I is approximately equal to 8 megohms. The voltage V_{o2} (at time t_h) is given by (Fig. 8)

$$V_{o2} = V_{o1}\, e^{\frac{-t}{R_I C}}. \tag{9}$$

In this system,

$$V_{o2} = 99.981\%\ V_{o1}.$$

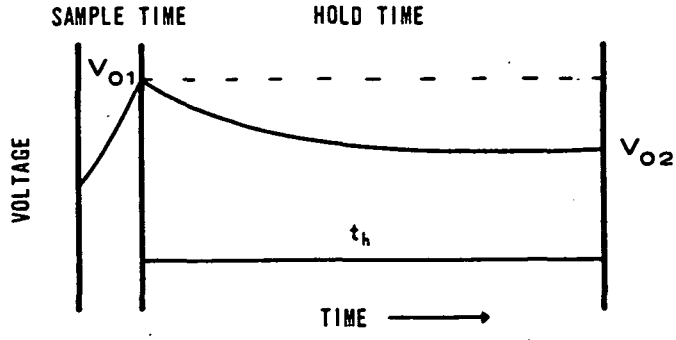


FIGURE 8. VOLTAGE DECAY DURING HOLD TIME

V_{O2} decays .019 percent at the end of the A/D conversion of all 32 channels to digital information. This gives an amplitude inaccuracy because of the A/D converter and sample-and-hold of .081 percent which is within the desired total system accuracy.

The A/D converter is required to produce a word every 48 μs; therefore, the word rate is 20.8 kHz. This conversion rate is generated by a crystal-controlled clock which also furnishes all systems timing to a high degree of accuracy.

The remaining unit to be described is the acquisition memory located in the format logic block of Figure 5. It is required because the data are generated continuously, but can only be recorded intermittently since the memory size of the data reduction computer restricts the amount of data it can accept at any given time.

Therefore, to avoid the loss of any test data, a magnetic core memory is used to store the data during the writing of a 1.907 cm blank space, l_g, on the tape which the computer recognizes as a stopping point when reading in. The number of tape characters to be stored, C_g, is given by

$$C_{gmax} = D\, l_g = 218.7 \times 1.907 = 417 \text{ characters.} \quad (10)$$

Since every third tape character is a zero and not written out of acquisition memory, it is necessary to store only two thirds of this value, which gives the number of acquisition words to be stored:

$$C_g = 2/3\ (417) = 278 \text{ characters (7 bit).} \quad (11)$$

The acquisition memory size for this system is 512 x 7 which is more than adequate. With a load rate of 32 acquisition words per sample and 34 word times for unloading, the memory gains two locations each sample time; 139 locations are required to store the 278 characters (equation 11). After 70 samples, 140 locations are available to store data while the tape gap is written. This permits the memory to be operated in a synchronous sequential interlace cycle, resulting in reliable operation and allowing use of a smaller memory stack size. (This operation was preferred to asynchronous operation, in which each record is written in memory and then transferred to tape, requiring that at least 2 records be stored in memory.) The necessary record length is then 70 x 32 = 2240 acquisition words.

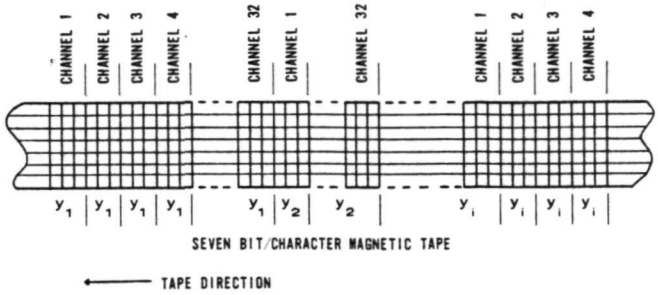


FIGURE 9. TAPE FORMAT

The tape format is shown in Figure 9. Each point per function is represented by three tape characters. The 32 sampled points, one per function, are written serially across the tape until 6720 characters (2240 acquisition words) are written. Then the record gap is written and the process repeats until the "stop recording" signal is given.

A photograph of the acquisition unit is shown in Figure 10.

FIGURE 10. ACQUISITION UNIT

REDUCTION UNIT

The reduction unit consists of a general purpose digital computer and a computer program written for handling and reducing the acquired data. Operating under control of the program, the computer determines the frequency from the digitized reference voltage recorded on an information channel and then calculates the first-harmonic phase-amplitude relationships at every test frequency for each channel from the digitized analog information.

To perform this calculation, it is necessary to fit a finite trigonometric sum to these recorded values (x_i, y_i). A method developed to accomplish this is based on the Fischer-Hinnen method of graphically determining the Fourier coefficient of a periodic function (Ref. 4).

The first harmonic Fourier coefficients are approximated by

$$A_1 = \frac{2}{m} \sum_{k=1}^{k=m} y_k \sin \left[(k - 1/2) \frac{2\pi}{m} \right] \qquad (12)$$

and

$$B_1 = \frac{2}{m} \sum_{k=1}^{k=m} y_k \cos \left[(k - 1/2) \frac{2\pi}{m} \right] \qquad (13)$$

where m = number of equal divisions of the period T; the precision in determining A_1 and B_1 increases by making m large. A_1 is twice the average height of the ordinates, which are obtained by multiplying y_1 by the sine of the angle measured from the origin to the ordinate y_1, multiplying y_2 by the sine of the angle measured from the origin to the ordinate y_2, etc. (Ref. 5).

Determining the Fourier coefficient in this manner requires that the origin be known or chosen. In this system the oscillator voltage that produces the shaker force is recorded on an information channel. The reduction computer program then searches this channel for the first positive slope zero crossover point. The magnetic tape record at which this occurs is the reference for all information channels. The correlation between the analog voltages and their sampled (later digitized) form is sketched in Figure 11.

The probability of sampling on the zero crossover is remote in a continuously sampling system. However, the reduction computer is programed to linearly

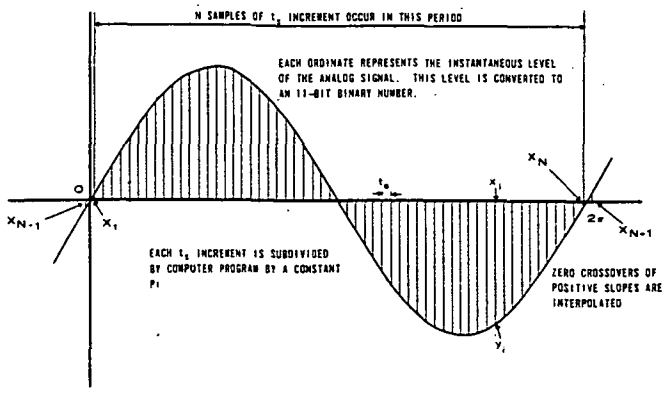

N SAMPLES OF t_s INCREMENT OCCUR IN THIS PERIOD
EACH ORDINATE REPRESENTS THE INSTANTANEOUS LEVEL OF THE ANALOG SIGNAL. THIS LEVEL IS CONVERTED TO AN 11-BIT BINARY NUMBER.
x_N
x_{N-1}
x_1
t_s
2π
EACH t_s INCREMENT IS SUBDIVIDED BY COMPUTER PROGRAM BY A CONSTANT P_i
ZERO CROSSOVERS OF POSITIVE SLOPES ARE INTERPOLATED
y_i

FIGURE 11. CORRELATION BETWEEN A TYPICAL ANALOG SIGNAL AND ITS SAMPLED FORM

interpolate between samples to position the reference as closely as system accuracies demand.

For this system, an m of 1200 was chosen to yield compatible accuracies. Since m is fixed, an interpolation constant p_i is calculated to the nearest whole number to give an m of 1200 for the period T. Thus, m can be expressed as

$$m = N p_i + q - \ell \qquad (14)$$

where N = number of digitized values between the zero crossover points of successive positive slopes of the reference voltage (oscillator), q is the number of subdivisions from x_1 back to the origin, and ℓ is the number of subdivisions from x_{N+1} to the zero crossover. Therefore,

$$p_i = \frac{m - q + \ell}{N} \qquad (15)$$

to the nearest whole number (Fig. 12).

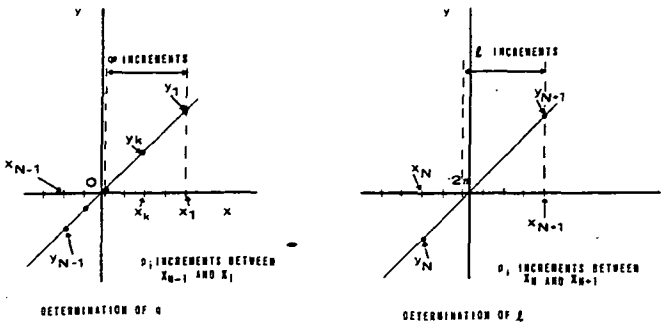


FIGURE 12. POSITIVE SLOPE ZERO CROSSOVER POSITIONING OF FOURIER REFERENCE BY LINEAR INTERPOLATION

The advantage of this interpolation is quite pronounced at higher frequencies. Without interpolation, the phase distortion introduced by positioning the origin on a sample point could vary by an entire sample time t_s. At 30 Hz (maximum frequency), the phase angle resolution α is

$$\alpha = \frac{2\pi}{m_o} = \frac{2\pi}{(SR)T} \simeq .3 \text{ radians} . \quad (16)$$

With interpolation, m = 1200,

$$\alpha = \frac{2\pi}{1200} = .0052 \text{ radians}$$

where $\alpha_{\text{allowable}} = .01748$ radians = 1.0 degrees.

y_k in equations 12 and 13 now becomes

$$y_k = y_{i-1} + \left(\frac{y_i - y_{i-1}}{P_i}\right) n \qquad n = 1, 2, 3, \ldots p_i. \quad (17)$$

The Fourier coefficient expressions now include both interpolated and measured values.

The actual frequency is computed by

$$f = \frac{1}{T} = \frac{1}{m \cdot \Delta x} = \frac{p_i}{m(x_i - x_{i-1})} = \frac{p_i}{mt_s} \quad (18)$$

where

$$\Delta x = \frac{(x_i - x_{i-1})}{p_i} = \frac{t_s}{p_i}.$$

The computation of the frequency and the coefficients A_1 and B_1 yields the response function,

$$y = A_1 \sin \omega_1 t + B_1 \cos \omega_1 t \quad (19)$$

where

$$\omega = 2\pi f = \frac{2\pi p_i}{mt_s}. \quad (20)$$

Equation 19 may also be written as

$$y = C_1 \sin (\omega_1 t + \phi) \quad (21)$$

where

$$C_1 = \sqrt{A_1^2 + B_1^2} \quad (22)$$

represents the amplitude and

$$\phi = \tan^{-1} \frac{B_1}{A_1} \quad (23)$$

represents the phase of the desired elastic-body response function.

CONCLUSIONS

Application of digital techniques to the problem of acquiring and reducing dynamic test data yields accurate structural gain and phase characteristics rapidly. It allows the control system designer to concentrate his efforts on the design aspects by freeing him from the time consuming chore of manually reducing the data. Errors inherent in manual data reduction are eliminated and use of the Fourier analysis reduces the effect of wave distortions. Accuracies of 0.15 percent of full scale amplitude and cross-channel phase correlation of less than 1.0 degree referenced to 30 Hz are realized with this system.

On SA-5 and SA-6, an average day of testing yielded 200 runs, where a run corresponds to the data collection for one frequency setting. An average of 350 samples per cycle for five cycles of acquired data for one day gave 11,200,000 data points for 32 channels requiring 33,600,000 tape characters to be reduced by the computer. These data are reduced and available for the control systems designer the day following the test.

Automating the test sequence to reduce human operation errors would improve the system. The test might be conducted from paper or magnetic tapes prepared prior to testing. The operator tape could then be used to digitally set the shaker driving oscillator. This would insure short term oscillator stability by closed loop digital control and allow setting of finer frequency increments about resonance peaks.

REFERENCES

1. Gaines, W. M., and Fischer, P. P., "Tentative Recommendations for the Definition and Testing of the Dynamic Performance of Analog to Digital Converters," AIEE Transactions on Communications and Electronics, Paper No. 61-15.

2. Ledley, R. S., Digital Computer and Control Engineering, McGraw-Hill Book Co., Inc., 1960.

3. Grabbe, Eugene, "et al," Handbook of Automation, Computation, and Control, Vol. 2, John Wiley and Sons, Inc., 1959.

4. Lawrence, R. R., Principles of Alternating Currents, 2nd Edition, McGraw-Hill Book Co., Inc., 1935.

5. Tang, K. Y., Alternating Current Circuits, 4th Edition, International Textbook Co., 1955.

H. SIGNAL CONDITIONING CIRCUIT FOR DISPLAYING PITCH, ROLL, AND YAW ON AN OSCILLOSCOPE

by Charles P. Elms

The problem of simultaneously displaying pitch, roll, and yaw attitude error information on an oscilloscope is treated by the programing of an analog computer and a unique switching circuit.

As a result of this study, a signal conditioning package for displaying the attitude signals on any laboratory oscilloscope was designed and produced.

NOMENCLATURE

ϕ_p	pitch attitude error
ϕ_Y	yaw attitude error
ϕ_R	roll attitude error
$\phi_{p'}$	conditioned pitch attitude error
$\phi_{Y'}$	conditioned yaw attitude error
S(t)	modulated sine component of ϕ_R
C(t)	modulated cosine component of ϕ_R
E sin ωt	sweep generator output

INTRODUCTION

A limited manual control simulation capability has recently been established to support the joint Marshall Space Flight Center-Ames Research Center investigation on pilot participation in guidance and control of the Saturn V launch vehicle. This article describes a technique used in this simulation for the simultaneous display of several variables. During actual flight, the control information is usually presented on the flight director attitude indicator (Ref. 1).

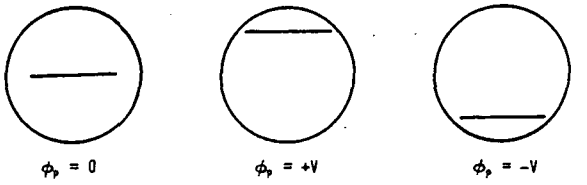
Two methods of conditioning the attitude signals for display on an oscilloscope were considered. The first consisted of fixing two perpendicular lines on the scope face to move in three degrees of freedom; vertical, lateral, and rotational. The second consisted of sweeping three separate lines that move independently of each other with one moving vertically, one laterally, and one rotationally. The second method was selected because of the complexity of the circuitry involved in implementing the fixed perpendicular lines.

This method consists of time-sharing the horizontal and vertical inputs of the oscilloscope. The traces would have to be swept one at a time at such a repetition rate that all would seem to appear on the screen at the same time. The time-sharing idea necessitates the use of switching circuits for timing and analog type circuits for conditioning of the attitude signals.

SYSTEM DESCRIPTION

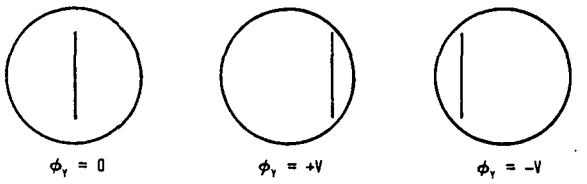
The system contains a sweep generator, a clock generator (Ref. 2), a three-stage ring counter (Ref. 3), a buffer, six gates, a diode function sine-cosine circuit (Ref. 4), two quarter-square multiplier circuits (Ref. 4), and eleven operational amplifiers (Ref. 5). The analog section of the system is programed on an analog computer and the switching is performed by a specially-built switching circuit.

Figure 1 depicts the traces on the oscilloscope that result from the three independent inputs. Figure 1(a) shows the conditioned pitch signal $\phi_{p'}$ entering the vertical input and the sweep signal E sin ωt entering the horizontal input. The sweep signal is sweeping the horizontal line and the pitch signal is biasing this line up and down proportionally to ϕ_p. Figure 1(b) shows the conditioned yaw signal $\phi_{Y'}$ entering the horizontal input and the sweep signal E sin ωt entering the vertical input. The sweep signal is sweeping the vertical line and the yaw signal is biasing this line left or right proportionally to ϕ_Y. Figure 1(c) shows a modulated cosine component of the roll signal entering the vertical input and a modulated sine component of roll entering the horizontal input. In this case, the oscilloscope is taking the vector sum and displaying the resultant as the roll attitude. This resultant is of constant magnitude because the sine component of ϕ_R



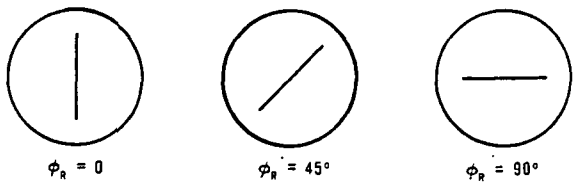
VERTICAL INPUT = ϕ_p', HORIZONTAL INPUT = E sin ωt

(a) PITCH DISPLAY - INTERVAL 1



VERTICAL INPUT = E sin ωt, HORIZONTAL INPUT = ϕ_Y'

(b) YAW DISPLAY - INTERVAL 2



VERTICAL INPUT = E sin ωt cos ϕ_R, HORIZONTAL INPUT = E sin ωt sin ϕ_R

(c) ROLL DISPLAY - INTERVAL 3

FIGURE 1. TIME-SHARED OSCILLOSCOPE INPUTS

varies directly with the cosine. Both inputs are in phase. To prove that the magnitude is always constant,

$$R(\phi_R) = E \sin \omega t (\sin \phi_R + j \cos \phi_R) = E \sin \omega t \, e^{j\phi R}.$$

An overall system block diagram is given in Figure 2. Pitch and yaw circuits are identical; therefore, only the pitch circuit is described in detail.

The pitch attitude signal ϕ_p is first attenuated and summed with a bias voltage V_p to fit its gate circuit.

$$\phi_{p'} = A_p \phi_p + V_p.$$

ϕ_p' is then gated, along with E sin ωt, through gates 1 and 2 during interval 1 and directed to the oscilloscope inputs, thus displaying a straight horizontal line that moves vertically as ϕ_p is varied.

The yaw attitude is conditioned the same as ϕ_p.

$$\phi_{Y'} = A_Y \phi_Y + V_Y.$$

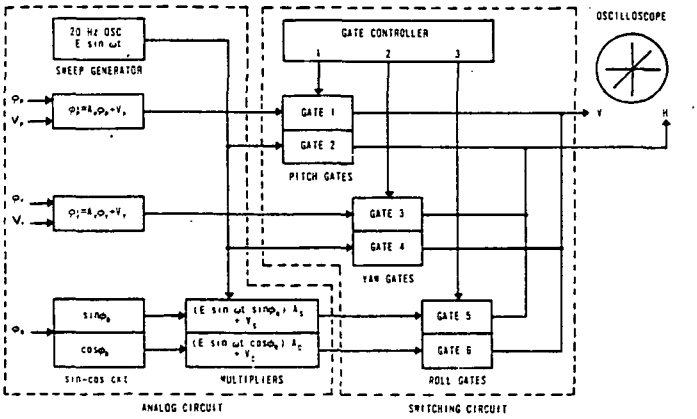


FIGURE 2. SYSTEM BLOCK DIAGRAM

In this case $\phi_{Y'}$ is gated, along with E sin ωt, through gates 3 and 4 during interval 2, to the oscilloscope inputs, thus displaying a vertical line that moves left and right as ϕ_Y is varied.

The roll attitude signal ϕ_R is first transformed into its sine and cosine components through a sine-cosine circuit. The outputs of the sine-cosine circuit are -sin ϕ_R and -cos ϕ_R. These components are then modulated with E sin ωt by multiplication with the sweep signal through two fixed diode function multipliers. These products are then attenuated and summed with a bias voltage to fit their particular gates.

$$S(t) = A_S E \sin \omega t \sin \phi_R + V_S$$

$$C(t) = A_C E \sin \omega t \cos \phi_R + V_C.$$

The S(t) and C(t) signals are gated through gates 5 and 6 to the oscilloscope inputs, thus displaying a straight line that can rotate clockwise for positive roll and counterclockwise for negative roll. Zero degrees roll is shown in Figure 1(c). Figure 3 shows the

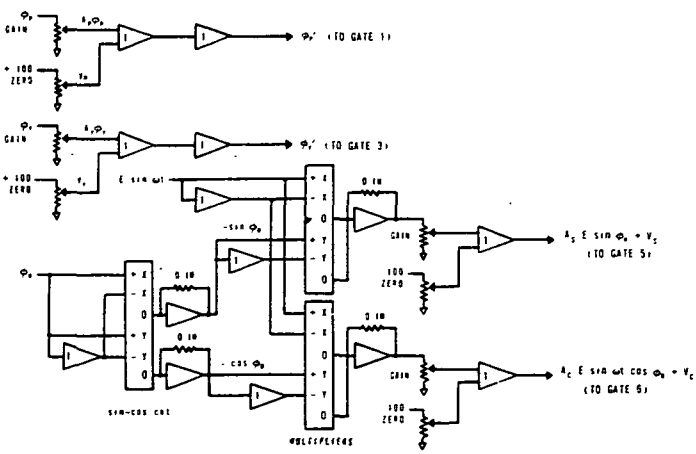
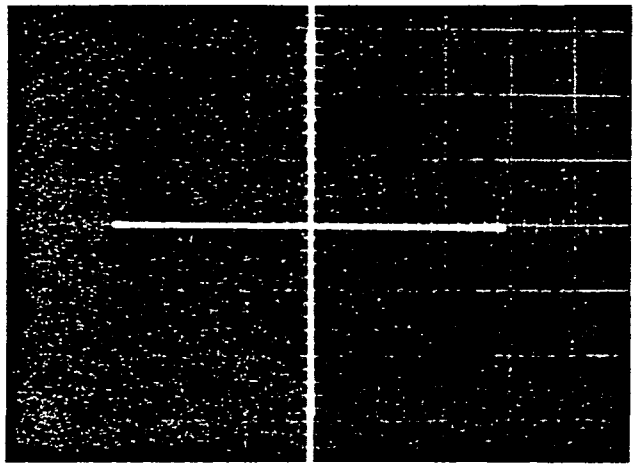
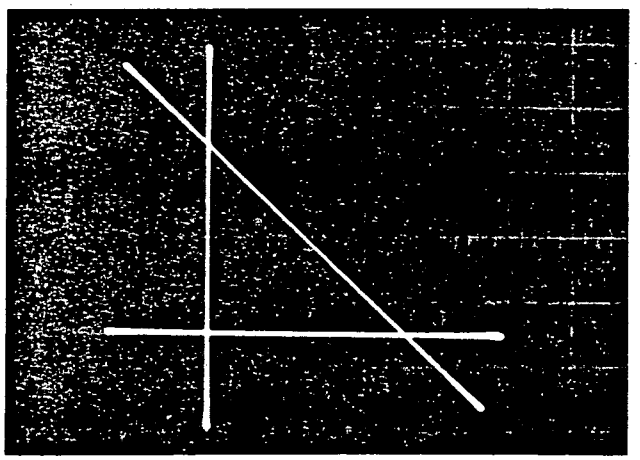


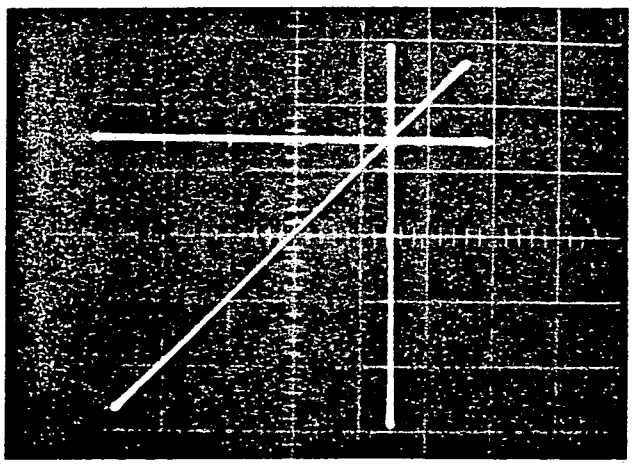
FIGURE 3. ANALOG COMPUTER PROGRAM



(a) $\phi_p = 0$, $\phi_r = 0$, $\phi_a = 0$



(b) $\phi_p = -5°$, $\phi_r = -5°$, $\phi_a = -45°$



(c) $\phi_p = +5°$, $\phi_r = +5°$, $\phi_a = +45°$

FIGURE 4. ACTUAL PHOTOGRAPHS OF ATTITUDE DISPLAYS

program as patched on the analog computer. Three different pictures taken of the oscilloscope for different pitch, roll, and yaw values are shown in Figure 4.

SWITCHING CIRCUIT

The switching circuit contains the clock generator, ring counter, buffer, and six gates. This progressive type of switching circuit closes gates 1 and 2 during interval 1, gates 3 and 4 during interval 2, and gates 5 and 6 during interval 3. A relationship between the clock generator output and ring counter outputs is given in Figure 5. When voltage to the circuit is first applied, the first stage of the ring counter is turned on. Pulse 2 cuts off stage 1 which turns on stage 2. Pulse 3 cuts off stage 2 which turns on stage 3. Pulse 1 cuts off stage 3 which turns on stage 1.

The outputs of the ring counter stages are negative pulses denoted as intervals 1, 2, and 3. These outputs are fed through a buffer for impedance matching and then directed to their particular gates.

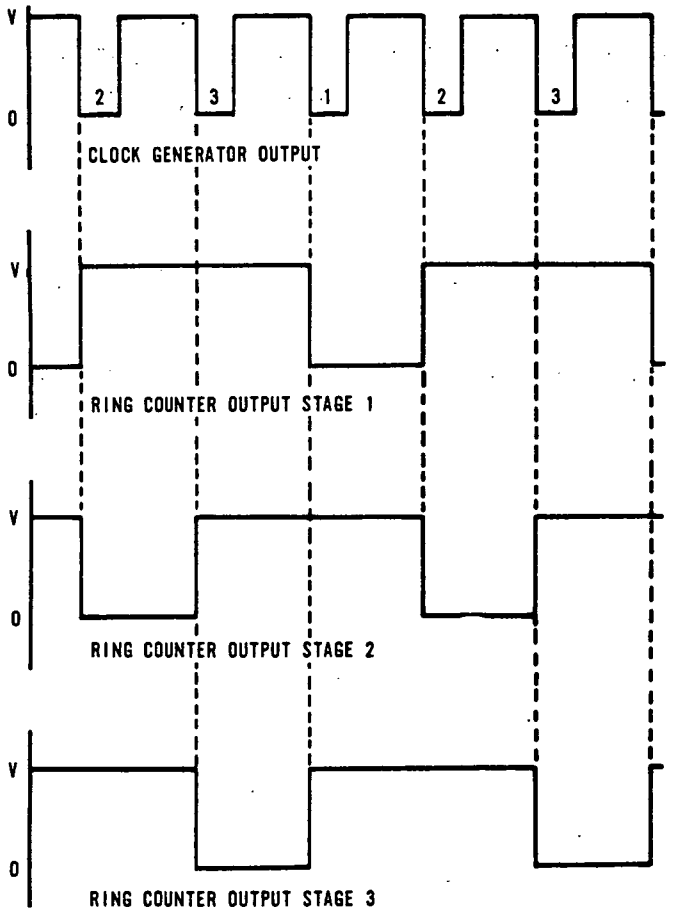


FIGURE 5. RELATIONSHIP BETWEEN CLOCK TIMING PULSES AND RING COUNTER OUTPUT PULSES

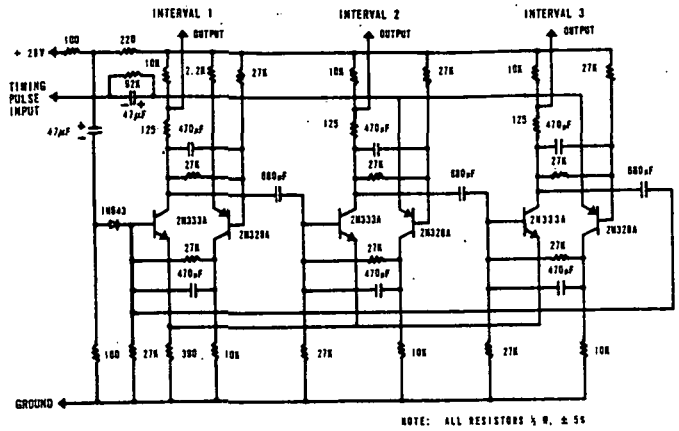


FIGURE 6. THREE-STAGE RING COUNTER

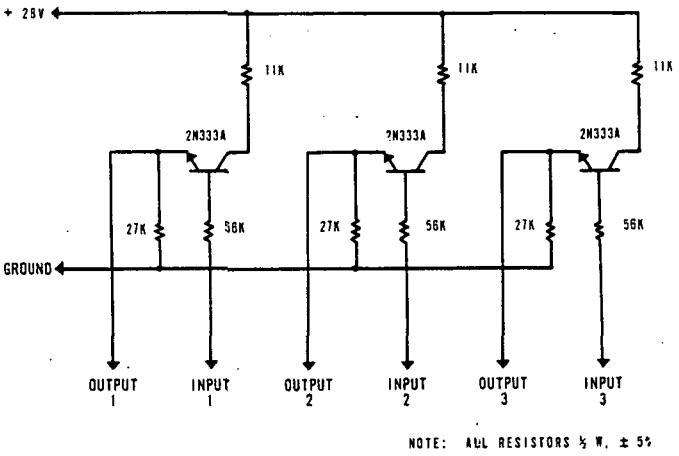


FIGURE 7. BUFFER CIRCUIT

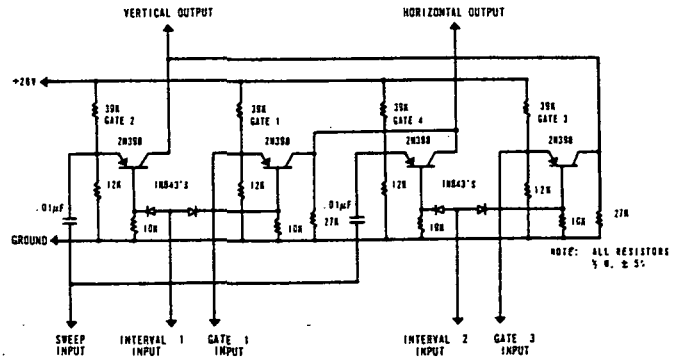


FIGURE 8. PITCH AND YAW GATES

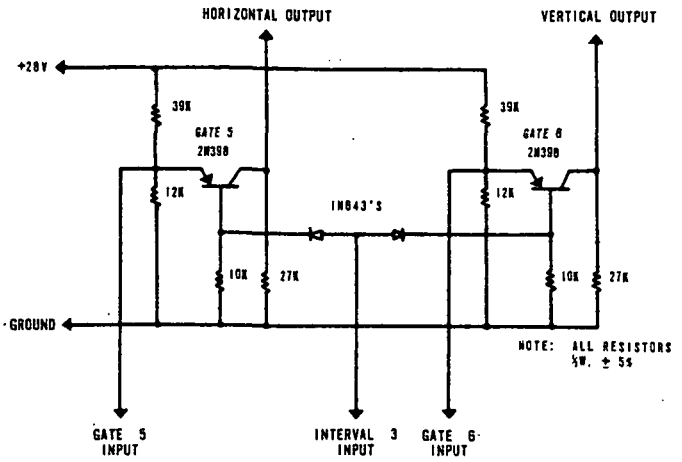


FIGURE 9. ROLL GATE

Figures 6, 7, 8, and 9 are schematics of the ring counter, buffer, pitch and yaw gates, and roll gates, respectively. The circuits are not discussed since they are self-descriptive. Caution must be taken in entering the switching pulses where two gates are switched by the same pulse. Diodes are used to block any crosstalk that might leak from one gate to the other.

CONCLUSIONS

This circuit has been installed and is functioning satisfactorily; it has operated for some 700 hours without maintenance other than periodic recalibration.

The main advantage is that this system can be used with any laboratory oscilloscope. Other advantages are the simplicity of individual circuits, the use of standard analog equipment, and its reliability for long periods of time.

Consideration has also been given to displaying not only pitch, roll, and yaw attitude error information, but also normal and lateral accelerations by expanding the existing system to include the method of fixed perpendicular lines. Thus, the three attitudes could be displayed by the fixed perpendicular lines and the two accelerations by two other independent perpendicular lines, as pitch and yaw are presented now. Research and development work is presently being done to implement this expansion.

REFERENCES

1. Flight Director Attitude Indicator, Data Sheet APD-147 Minneapolis Honeywell Reg. Co., May 1963.

2. HF3 Clock Generator and Multivibrator, Packard Bell Computer Corp. Data Sheet.

3. Hand Book of Selected Semiconductor Circuits, prepared by Bureau of Ships, Dept. of the Navy, April 1, 1962.

4. Multipliers and Fixed Diode Function Generators, Applied Dynamics, Inc., Handbook AD-2-64PBC, Computer Modules Index No. 12, October 1963.

5. Amplifier Assemblies, Applied Dynamics, Inc., Handbook AD-2-64PBC, Computer Modules Index No. 9, October 1963.

1. AN ANGULAR POSITION CONVERTER

by Leslie C. Holt

The angular position converter is used to determine the angular position of the command attitude modules located in the Saturn guidance signal processor. In describing this method, emphasis has been placed on obtaining attitude information in analog form on existing guidance system equipment as used in flight simulation and hardware checkout problems.

A breadboard model converter has been designed and built to utilize the existing Optisyn output signals. The generated attitude signal in analog form was found to be sufficiently accurate for test purposes.

INTRODUCTION

The Saturn guidance signal processor (GSP) is the input-output unit for the digital guidance computer and other vehicle systems (Ref. 1). The GSP is the interconnecting unit of the guidance system where interface changes, caused by mission requirements, are incorporated. In guidance simulation and hardware checkout, this interface consists of an angular position converter as shown in Figure 1.

The angular position converter is used to accurately determine the angular position of the command attitude modules (located in the GSP) for flight simulation and hardware checkout. Since the existing guidance system equipment used in flight simulation does not provide the module position in an acceptable analog form, and since the equipment cannot easily be modified to provide the desired data, the available signals must be converted to a voltage suitable for analog computer use.

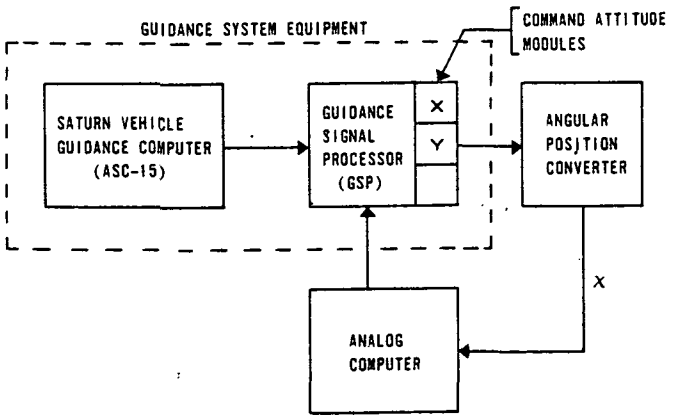


FIGURE 1. BLOCK DIAGRAM OF SIMULATED GUIDANCE SYSTEM

The dynamic command attitude modules are motor driven to the required orientation (computed by the flight computer) of the vehicle. As the modules are rotated, the shaft position encoders (Optisyns) generate signals that are proportional to the angle of rotation.

The Optisyn output signals are two sine waves displaced by 90 degrees. The signals are sent to the angular position converter where they are shaped into square waves. From the square waves, the coder recognizes a change in the bit configuration and an increment is added to or subtracted from the accumulated value held in the counter register. The register contents are then converted to an analog voltage and sent to the analog computer. The analog computer, simulating flight mechanics and control dynamics, generates velocity coordinates that are sent back to the GSP and closes the simulator loop.

NOMENCLATURE

Q, R	Optisyn output signals (sine waves displaced by 90 degrees)
Q', R'	shaper driver output signals (square waves in reflected binary)
Q'', R''	Gray to binary output signals (serial binary form)
X	module position in analog form
θa	actual shaft position, degrees
θm	measured shaft position, degrees

OPTISYN SHAFT POSITION ENCODER

The Optisyn is an electromechanical device employing an optical technique to convert shaft motion into two sine waveforms that can be counted digitally. The Optisyn uses two transparent disks having photo patterns of alternately opaque and clear sectors. One disk has one more opaque and transparent sector than the other, resulting in a light interference pattern.

The two disks are mounted concentric to the Optisyn axis of rotation, with one disk mounted on the input shaft. In operation, the lamps transmit light through the interfering disk patterns to the silicon cells, providing the relative light intensity pattern shown in Figure 2. Maximum light will be transmitted at one point on the disk circumference; at the same time, minimum light will be 180 degrees away (Ref. 2).

The Optisyn output frequency is 150 electrical cycles per 360 degrees of mechanical shaft rotation. The resolution (number of output steps) then is four times the number of electrical cycles per mechanical revolution of the Optisyn shaft. Since the Optisyn shaft is connected to the module shaft through a 12 to 1 ratio gear train, one output voltage step is equivalent to 0.05 degree of module rotation.

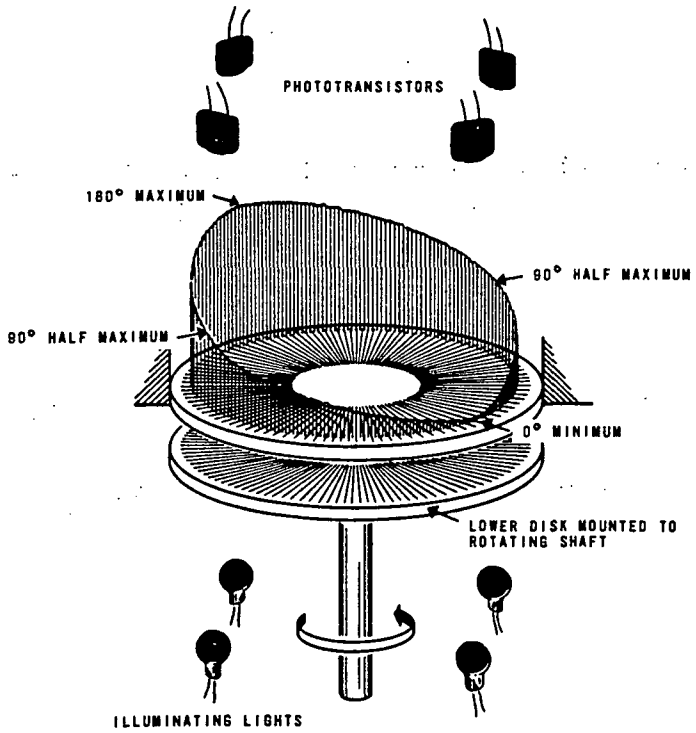


FIGURE 2. OPTISYN SHAFT ENCODER (taken from a brochure by Dynamics Research Corp.)

OPTISYN CONVERTER

The Optisyn output signals are two sine waves, 90 degrees apart in phase and proportional to the change in each command angle. The signals, identified as Q and R, are shown in Figure 3 at a changeover where

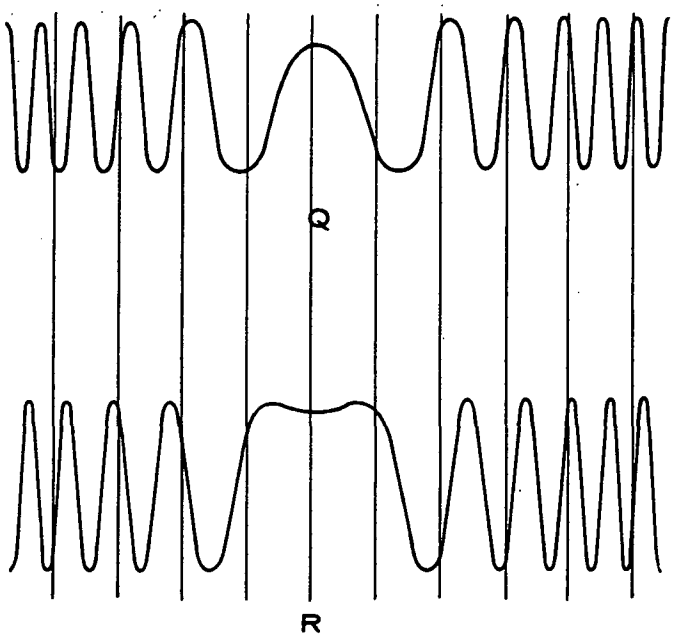


FIGURE 3. OPTISYN OUTPUTS

Q changes from lag to lead R. This change in phase relationship represents a change in direction of rotation of the command attitude module axis being sensed.

The Optisyn output signals are fed into shaper-driver circuits, as shown in Figure 4, to form a two-bit reflected binary number, each bit of which appears on a separate line (Ref. 3). The reflected binary signals Q' and R' are next converted to true

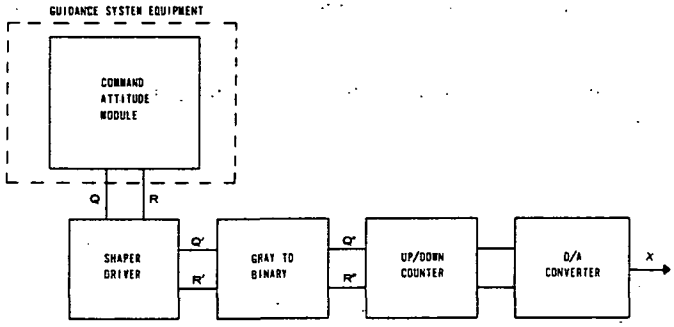


FIGURE 4. CONVERTER BLOCK DIAGRAM

binary Q'' and R'' as shown in Figure 5. The true binary form is suitable for use in a digital to analog (D/A) converter.

An up/down binary counter is used to hold 9 bits of the 11-bit angular position word and also serves as a register for the D/A converter. The two serial

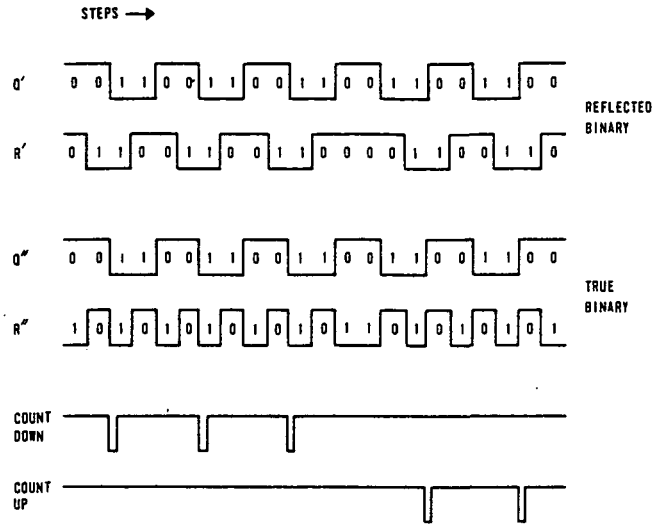


FIGURE 5. CONVERTER TIMING CHART

binary signals, R'' and Q'', respectively (Fig. 5), are used as the two least significant bits of the D/A register-counter. The decision to increase or decrease the 9-bit up/down counter, depending on direction of module rotation, is made from the reflected binary (Gray) word Q' R' as shown in Figure 5.

STEPS	REFLECTED BINARY (GRAY)		TRUE BINARY	
	Q′	R′	Q″	R″
1	0	0	0	1
2	0	1	0	0
3	1	1	1	1
4	1	0	1	0
5	0	0	0	1
6	0	1	0	0
7	1	1	1	1
8	1	0	1	0
9	0	0	0	1

FIGURE 6. REFLECTED BINARY TO BINARY CONVERSION

In steps 1 to 5 of Figure 6, the shaft angle decreases one bit per step; in steps 5 to 1, the shaft angle increases one bit per step. In step 3, Q' R' and Q'' R'' are 11, which indicates a maximum 2-bit word. If the input (Q' R') is changed to the condition shown in step 2, Q' R' is changed to 01 (Q'' R'' is 00) which indicates a carry one or add one bit to the up/down counter contents. Likewise, when going from step 3 to step 4, Q' R' changes to 10 (Q'' R'' is 10) which indicates a borrow one or subtract one from the contents of the up/down counter. Each time the two low-order bits (Q'' R'') are changed from condition 11, the third-order bit (least significant bit of the up/down counter) is changed accordingly.

DIGITAL TO ANALOG CONVERTER

The 11-bit binary word, held in the counter-register, is transferred in parallel to the D/A converter. The input section of the D/A converter contains another 11-bit register where the data word is held in temporary storage for the convert signal. The convert signal is generated at both the rise and fall edges of the Q' and R' Optisyn output signals. Therefore, a change in level of the Q' or R' signals will generate a convert signal and, in turn, a change in the analog output voltage proportional to the Optisyn (Q' R') change.

The D/A output section is composed of a register ladder network. The output can have $2^{11} - 1$ or 2047 possible increments, depending on the 11-bit digital number placed in the converter. Since the ladder output has a maximum range of ± 5 volts, each ladder increment represents 4.9 mV.

The full 11-bit binary word represents 102.35 degrees (0.05 degree per bit) of shaft rotation. The bit size of the converter-register can be expanded to more bits with higher output resolution; however, higher resolution is not practical because of noise consideration in the present application.

ALTERNATE TECHNIQUE

An alternate method of obtaining an acceptable analog voltage for computer use is available in the existing guidance system equipment presently used for simulation. The output signal is generated in a shaft-mounted potentiometer for telemetry use in actual flight. Since the linearity of the potentiometer is 0.5 percent full scale, considerable error is introduced

into the simulation loop, causing unacceptable system response.

A comparison of the two methods of obtaining the desired signal is as follows.

Let θa = actual shaft position

θm = measured position

then for the potentiometer signal

$$|\theta m - \theta a| \leq 360\ (0.005)$$

$$|\theta m - \theta a| \leq 1.8 \text{ degrees}$$

Compared to the Optisyn converter,

$$|\theta m - \theta a| < 0.05 \text{ degree.}$$

This comparison proves that the converted Optisyn analog signal is more suitable for closed loop simulation than the potentiometer signal.

CONCLUSIONS

A breadboard model converter has been designed and built to utilize the existing Optisyn output signals, which may be converted to a single analog voltage suitable for use in an analog computer. The guidance simulation for accuracy checkout could not be performed without accuracy on the order of 0.1 degree. The angular position converter technique results in simulation accuracy of 0.05 degree, which is well within the 0.1 degree requirements. This device will also be useful for closed loop simulation problems in the advanced guidance system equipment.

Of the two methods presented, the angular position converter method is the most accurate. Since the available potentiometer linearity error is 0.5 percent, the shaft angle error can be 1.8 degrees as compared with 0.05 degree using the converter.

REFERENCES

1. Innes, D. J., "Functional Description of the Guidance Signal Processor (GSP-24)," Working Paper, Guidance and Control Systems Branch, Astrionics Division, MSFC.

2. Optisyn Shaft Position Encoder Brochure, Dynamics Research Corporation, Woburn, Mass.

3. The ASC-15 SVG Computer, International Business Machines Corporation, NASA Contract No. NAS 8-2565, January 3, 1963

II ELECTRICAL SYSTEMS...

A. GEOMETRIC ANALYSIS OF THE MEAN NORMALIZED POWER AVAILABLE FROM THE SOLAR CELLS ON THE MICROMETEOROID MEASUREMENT CAPSULES by Richard Boehme and Jimmy L. Miller

A method for analyzing those parameters incurred in the geometrical analysis of solar cell panel orientations with respect to the sun for a satellite in random tumble is presented. Also presented is the concept of mean normalized power available. This is obtained by averaging the power available from a sampling of possible orientations taken in a weighted manner.

The methods as applied to the solar cell panels of the micrometeoroid measurement satellite are included in the appendix.

INTRODUCTION

In the design of a solar cell power conversion system for satellite application, one of the more prominent considerations is that of cell array orientation with respect to the sun. Associated with this is the problem of shadowing of the cell arrays by either the earth or the satellite itself. Solar cell array orientation devices mounted on extended arms may be employed to better utilize the available solar energy, but at the cost of added complexity and lower overall reliability. The alternative is to rigidly mount solar cell arrays on the vehicle in a pattern that will produce the required power with a minimum of cells.

The number of cells required for each array must be determined by considering the summed output from all arrays at each possible vehicle orientation with respect to the sun. Those orientations which result in minimum power and those which result in maximum power are of prime interest. Total power available from the solar cells is determined by a consideration of total sun time as compared to total shadow time, which includes earth shadow and vehicle orientation shadow.

Because of the complex forces operating on a satellite, the tumble motions that often result, unless special attitude stabilization is employed, are extremely difficult to analyze. These motions are often described as random tumble. Slow oscillatory modes normally result after a period of time.

One approach for working with random tumble is discussed, and a method for finding the mean normalized power available from the solar cell panels on a typical satellite is presented.

RANDOM TUMBLE

To determine the mean power available during random tumble, one must first define random orientation in a workable manner. The word random implies that equal probability exists for each distinct orientation. To satisfy this criterion, it seems feasible to define a unit solar vector and to assign equal weighting to the elements of area generated on the surface of a sphere by this vector in randomly changing orientation. For simplicity of mathematics, the solar vector, S, is defined as a unit length vector oriented negatively to the direction of the solar flux velocity vector. In this concept, the sphere is generated by the head of the unit solar vector in random orientation with the tail fixed at the center of the sphere. In reality, however, the satellite is located at the center of the sphere and is in random orientation, while the solar vector is fixed outside the sphere (Fig. 1).

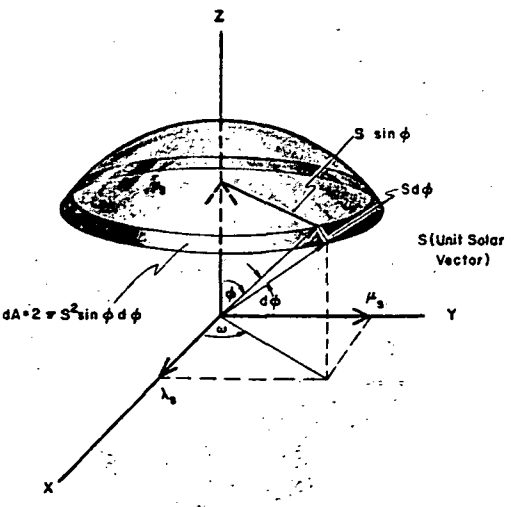


FIGURE 1. UNIT RADIUS SPHERE

Using standard orthogonal axes (X, Y, Z), the unit solar vector may be defined in terms of direction cosines (λ, μ, γ), as shown in Figure 1.

To orient S through all possible directions and describe the motion mathematically, it is necessary to systematically vary the orientation. Figure 1 shows that this variation is controlled by two parameters, ϕ and ω; ϕ is the angle the solar vector makes with the positive Z axis and ω is the angular displacement from the positive X axis of the projection of S onto the the X-Y plane.

The sphere with radius S is generated by allowing ω to rotate through multiples of 360 degrees, while ϕ goes from 0 to 180 degrees. Equal weighting for elements of spherical area is achieved by allowing ϕ to increase relative to ω in such a way as to allow the rate of area generation to be constant with respect to increase in ω.

The element of area is given by

$$dA = 2\pi S^2 \sin \phi \, d\phi. \tag{1}$$

Dividing by $d\omega$ produces an expression for the rate of area generation with respect to the increase in ω.

$$\frac{dA}{d\omega} = 2\pi S^2 \sin \phi \frac{d\phi}{d\omega}.$$

Setting $\frac{dA}{d\omega} = K$ and solving for ϕ relative to ω with the initial condition that when $\phi = 0$ degrees then $\omega = 0$ degrees produce

$$\int \sin \phi \, d\phi = \int \frac{K d\omega}{2\pi S^2}$$

$$-\cos \phi = \frac{K\omega}{2\pi S^2} + C_1. \tag{2}$$

The initial conditions may be substituted to solve for C_1.

$$-\cos 0^\circ = C_1$$

$$C_1 = -1.$$

Therefore,

$$-\cos \phi = \frac{K\omega}{2\pi S^2} - 1.$$

Since S is unity, it may be replaced and the relation rearranged.

$$\cos \phi = 1 - \frac{K\omega}{2\pi}. \tag{3}$$

By applying the trigonometric identity

$$\sin^2\phi + \cos^2\phi = 1$$

equation 3 may be used to obtain equation 4.

$$\sin \phi = \sqrt{\frac{K\omega}{\pi}\left(1 - \frac{K\omega}{4\pi}\right)}. \tag{4}$$

The constant, K, has the dimensions of area per degree and may be selected to give a desired number of rotations of ω during the variation of ϕ from 0 to 180 degrees. Selection of a smaller K gives a greater number of rotations and results in greater accuracy.

The direction cosines of S may be expressed in terms of ω by the following process:

$$\lambda_s = \sin \phi \cos \omega$$

$$\mu_s = \sin \phi \sin \omega$$

$$\gamma_s = \cos \phi.$$

Substituting equations 3 and 4,

$$\lambda_s = \sqrt{\frac{K\omega}{\pi}\left(1 - \frac{K\omega}{4\pi}\right)} \cos \omega \tag{5}$$

$$\mu_s = \sqrt{\frac{K\omega}{\pi}\left(1 - \frac{K\omega}{4\pi}\right)} \sin \omega \tag{6}$$

$$\gamma_s = 1 - \frac{K\omega}{2\pi}. \tag{7}$$

By the proper choice of K, these expressions for the solar vector direction cosines may be used with consecutive values of ω to give the desired number of rotations of the solar vector in sweeping out the imaginary sphere.

EFFECTIVE AREAS AND SHADOWING

The projected area of a solar cell panel is defined as the dot product of the normal to the face of that panel and the solar vector. The normal to the face of a solar cell panel may be denoted by N_i and its direction cosines by (λ_i, μ_i, γ_i). The projected area then is

$$A_i = N_i \cdot S$$

or

$$A_i = \lambda_i\lambda_s + \mu_i\mu_s + \gamma_i\gamma_s. \tag{8}$$

The total projected area for η solar cell panels is then given by

$$A_T = \sum_{i=1}^{\eta} A_i \qquad \text{for } A_i > 0. \tag{9}$$

The angle, ϵ, between the direction of the normal to the cell array surface and the direction of the solar vector S may be expressed by

$$\epsilon_i = \arccos(\lambda_i\lambda_s + \mu_i\mu_s + \gamma_i\gamma_s). \tag{10}$$

In this expression, ϵ is the angle of incidence. As ϵ increases, the cell response drops below the projected area because of reflectance from the cell and cover glass. A very close approximation to the response of solar cells shielded with cover glass is obtained by

$$\text{Response} = A_i - \rho\epsilon. \tag{11}$$

Figure 2 shows a comparison of projected area and cell response for cells shielded with 60 mil cover glass. For this particular glass, a value for ρ of 0.002 is appropriate.

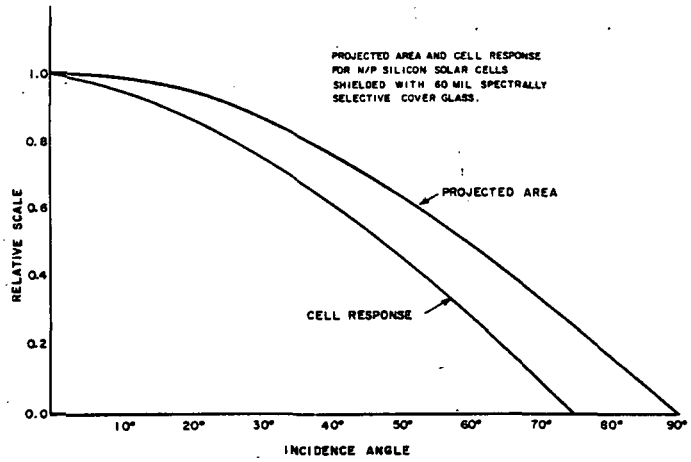


FIGURE 2. SOLAR CELL RESPONSE PROFILE

The calculation of a negative value for A_i in equation 8 means that ϵ is greater than 90 degrees and that the solar irradiance is not falling on the front surface of the panel. Therefore, in calculating the power obtainable from a solar cell panel (or for a number of panels as in equation 9), only positive values of A_i may be considered.

Graphical methods may be used to determine those combinations of ϕ and ω which result in a shadow from another portion of the vehicle being cast upon a particular solar cell panel. For a pessimistic appraisal of total power, those terms of A_i may be eliminated which represent combinations of ϕ and ω that produce a shadow upon any portion of that panel.

MEAN NORMALIZED POWER AVAILABLE FROM η PANELS

To compare the advantages of various solar cell panel arrangements on a satellite, a figure termed "mean normalized power" may be calculated. This term represents the mean power available from all panels during the variations of ϕ and ω required to sweep out an entire sphere about the vehicle. The mean power is normalized on the basis of the power available from one panel receiving light at normal incidence. In the event the panels are not of the same capacity, or differ in other respects, one may be chosen as the normalizing standard and the other terms of A_i weighted to allow for the deviation of panel capacity.

The total power is obtained by using equation 11 written for a summation of η panels.

$$P_T = \sum_{i=1}^{\eta} (A_i - \rho\,\epsilon_i). \tag{12}$$

By the appropriate substitutions into equation 12, the total power is given in terms of ω. These relations may be set into a computer and the program can be adjusted to compute total power for fixed increments of ω. The mean normalized power available is obtained by averaging the total power values over the entire sphere. By taking equal increments of ω, the relations derived will provide the desired distribution of power samplings over the sphere. This distribution will provide equal solid angle spacing for each sample calculated, thus implying equal probability for the existence of each distinct orientation.

APPENDIX

To demonstrate the methods presented, the solar cell panel deployment used on the micrometeoroid measurement capsule (MMC) is analyzed. The MMC will be attached to, and orbited with, the S-IV stage of a Saturn I block 2 launch vehicle. In orbit, the satellite will tumble randomly while collecting data on

the micrometeoroid hazard presented to space vehicles. The MMC uses four 6300-cell panels deployed around the satellite in the relationship of the sides of a regular tetrahedron.

Figure 3 shows a regular tetrahedron superimposed upon a set of rectangular coordinates in conformance with the MMC coordinate system. Since the

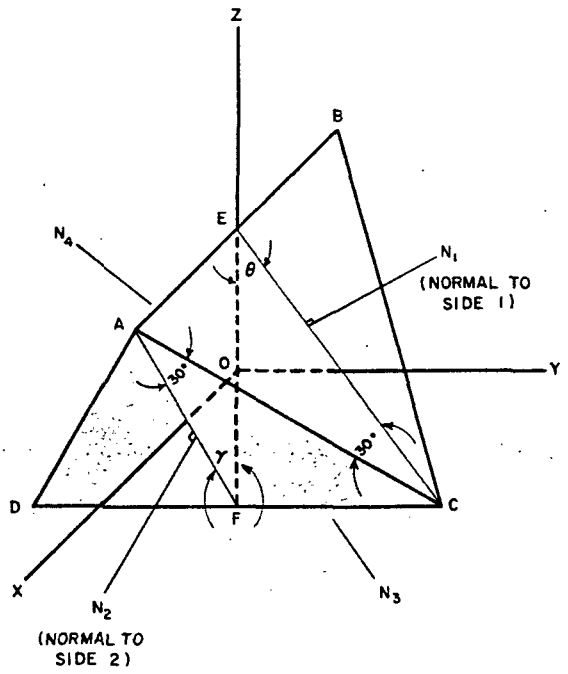


FIGURE 3. TETRAHEDRON

Z axis passes through the center of the tetrahedron, it bisects AB at point E and bisects CD at point F.

Therefore,

$\overline{AB} = \overline{AC} = \overline{CD}$

$\overline{CF} = 0.5\ \overline{CD} = 0.5\ \overline{AC}$

$\overline{CE} = \overline{AC} \cos 30°.$

Also since

$$\sin \theta = \frac{\overline{CF}}{\overline{CE}} = \frac{0.5\ \overline{AC}}{\overline{AC} \cos 30°}$$

$$\sin \theta = \frac{0.5}{\cos 30°}$$

$\theta = 35.3°.$

In a similar manner,

$\overline{AE} = 0.5\ \overline{AB} = 0.5\ \overline{AC}$

$\overline{AF} = \overline{AC} \cos 30°$

$$\sin \gamma = \frac{\overline{AE}}{\overline{AF}} = \frac{0.5\ \overline{AC}}{\overline{AC} \cos 30°}$$

$$\sin \gamma = \frac{0.5}{\cos 30°}$$

$\gamma = 35.3°.$

The normal to side 1 then makes an angle of 54.7 degrees (90.0 - 35.3) with respect to the positive Z axis and an angle of 35.3 degrees with respect to the positive Y axis. Since side 1 is parallel to the X axis, the normal to side 1 makes an angle of 90 degrees with respect to that axis. Similarly, the normal to side 2 makes an angle of 125.3 degrees (90.0 + 35.3) with respect to the positive Z axis, an angle of 35.3 degrees with respect to the positive X axis, and an angle of 90 degrees with respect to the Y axis.

The normals to the four sides may be expressed in terms of direction cosines as follows:

N_1 $\lambda_1 = \cos 90.0° = 0.0$

$\mu_1 = \cos 35.3° = 0.816$

$\gamma_1 = \cos 54.7° = 0.577$

N_2 $\lambda_2 = \cos 35.3° = 0.816$

$\mu_2 = \cos 90.0° = 0.0$

$\gamma_2 = \cos 125.3° = -0.577$

N_3 $\lambda_3 = \cos 144.7° = -0.816$

$\mu_3 = \cos 90.0° = 0.0$

$\gamma_3 = \cos 125.3° = -0.577$

N_4 $\lambda_4 = \cos 90.0° = 0.0$

$\mu_4 = \cos 144.7° = -0.816$

$\gamma_4 = \cos 54.7° = 0.577$

A computer was programed to substitute these direction cosines along with relations 5, 6, and 7 into equations 8 and 10.

$$A_i = \lambda_i\lambda_s + \mu_i\mu_s + \gamma_i\gamma_s \qquad (8)$$

$$\epsilon_i = \arccos(\lambda_i\lambda_s + \mu_i\mu_s + \gamma_i\gamma_s). \qquad (10)$$

A K value of $\frac{4\pi}{3600}$ was used and calculations were made for each 10-degree increment of ω. Printouts were made according to the following:

$$P_i = A_i - \rho\epsilon_i \qquad \text{for } i = 1, 2, 3, 4.$$

The total power summing and the elimination of terms because of shadowing were manually performed as was the averaging of terms to obtain the mean normalized power figure. Since the cells were covered with 60 mil glass, a value of 0.002 was used for ρ.

The computer analysis shows that the mean normalized power available is 0.799 without shadow consideration and 0.774 with shadow consideration.

Figures 4A and 4B show a plot of total power versus ω. The K value of $\frac{4\pi}{3600}$ produced ten revolutions of ω for a variation of ϕ from 0 to 180 degrees. The abscissa is divided into three scales showing ϕ, $\omega <$ 360 degrees, and ω total. On the ordinate, the total power is shown on a relative scale as a percentage of the power available from one panel alone at normal incidence.

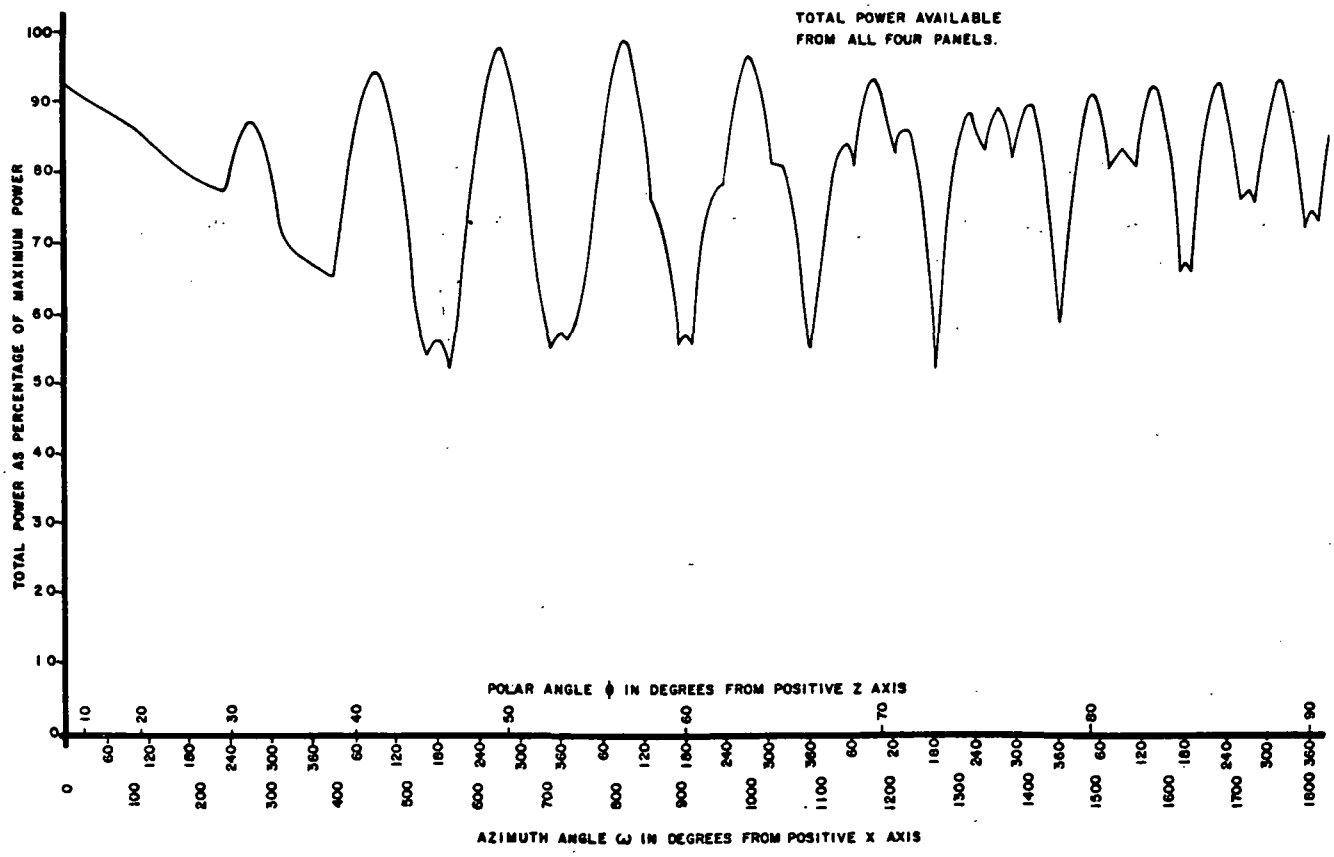


FIGURE 4A. TOTAL POWER AVAILABLE

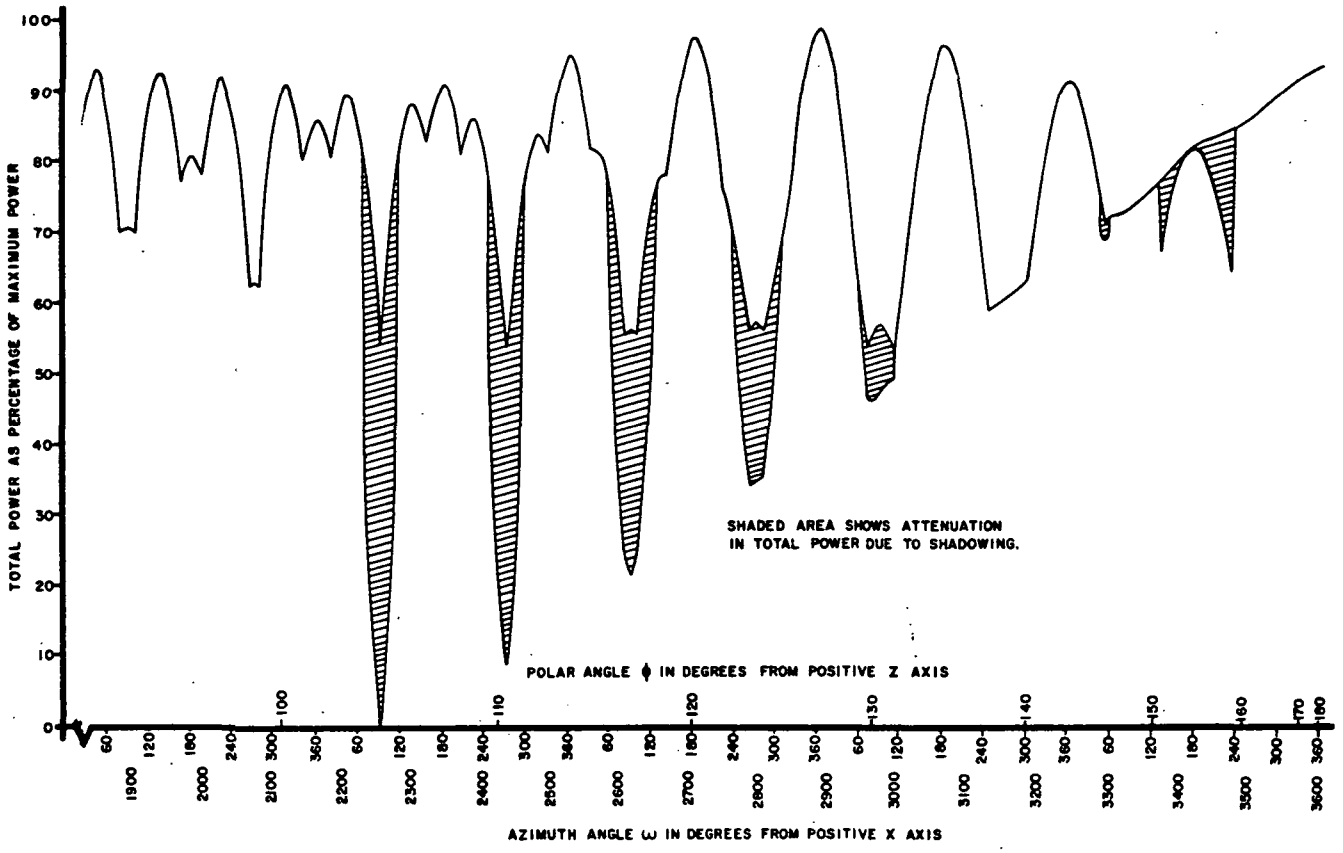

TOTAL POWER AS PERCENTAGE OF MAXIMUM POWER
SHADED AREA SHOWS ATTENUATION
IN TOTAL POWER DUE TO SHADOWING.
POLAR ANGLE ϕ IN DEGREES FROM POSITIVE Z AXIS
AZIMUTH ANGLE ω IN DEGREES FROM POSITIVE X AXIS

FIGURE 48. TOTAL POWER AVAILABLE

In certain orientations, it is possible for the MMC detector wing or the S-IV stage to cast a shadow upon a solar panel. Fortunately, this shadowing occurs at high angles of incidence where the power available from that panel is small. The shaded area shows the result of attenuation in total power caused by shadows.

This analysis of the MMC power collection solar cell system gave very high confidence in providing an adequate amount of power.

B. FUEL CELL POWER FOR SATURN INSTRUMENT UNIT

by Richard Boehme and Lance Conway

A hydrogen-oxygen fuel cell system that could be used to furnish electrical power for space vehicles is discussed. Fuel cell chemistry is covered only to the extent of helping the reader understand the basic technical aspects of fuel cell operation. Physical parameters of the fuel cell and the prospective integration scheme into a vehicle system are covered in enough detail to give an understanding of the fuel cell power supply.

INTRODUCTION

The electrical power and time requirements of space vehicles are increasing to such an extent that conventional battery systems will restrict the scope of missions. To provide adequate power for these missions, a fuel cell system is desirable because of its higher performance capabilities.

For the past several years, Marshall Space Flight Center has conducted research and development under an Office of Advanced Research and Technology (OART) program. The purpose of the program is to advance the state of the art and to establish technology necessary for a fuel cell system which could meet the wide range of applications and stringent requirements imposed on electrical equipment for space vehicles. The fuel cell system described reflects the results of work performed by the Allis-Chalmers Manufacturing Company for MSFC under this program.

The Allis-Chalmers fuel cell is a hydrogen-oxygen, medium-low temperature, low-pressure system of the capillary-held electrolyte type. Compared to fuel cells of the solid electrolyte or ion-exchange membrane types, this system is considered the most desirable for space vehicle applications because of its superior performance, flexibility, and reliability.

Recent studies have shown that, for extended mission times, such a fuel cell affords the following important advantages over present battery systems.

1. Reduction of the power system weight by greatly improving the energy-to-weight ratio.

2. Reduction of equipment volume and mounting requirements.

3. Improvement of electrical performance and regulation, and allowance for larger load variations. The poor regulation of conventional batteries requires the use of power-using regulation equipment. This decreases system efficiency and increases the burden on the cooling system. A regulation tolerance of at least ± 10 percent is necessary for a battery system.

4. Improvement of flexibility to allow for large variations in operating time or power profiles without major equipment redesign.

5. Provision of onboard power during prelaunch operations, thereby eliminating requirements for critical power transfer from ground support equipment to the vehicle. This can reduce switching complexity and increase the reliability of distribution.

6. Provision of potable water for manned missions.

FUEL CELL CHEMISTRY

The fuel cell is an electrochemical converter in which the reactants, or fuel, and the reaction products are not stored within the case as they are in a battery. In the fuel cell, reactants are stored separately and fed continuously to the electrodes, with the reaction products being removed continuously. As long as the processes continue, the fuel cell produces electrical energy. The reactants in the Allis-Chalmers fuel cell are hydrogen and oxygen and the reaction byproducts

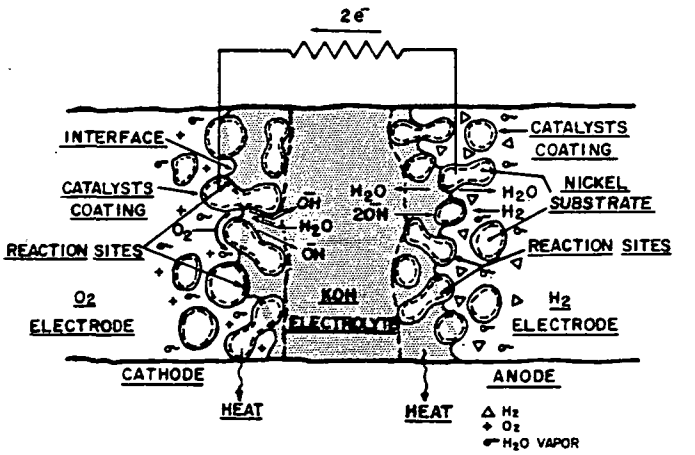


FIGURE 1. CAPILLARY FUEL CELL REACTION MECHANISMS

are heat and water (Fig. 1). Essentially, the simplified reactions are:

At the anode,

$$H_2 + 2\ OH^- \rightarrow 2\ H_2O + 2e^-$$

At the cathode,

$$\tfrac{1}{2}\ O_2 + H_2O + 2e^- \rightarrow 2OH^-.$$

Overall reaction,

$$H_2 + \tfrac{1}{2}\ O_2 \rightarrow H_2O + \text{heat} + \text{electrical energy}.$$

FUEL CELL CONSTRUCTION

The cell consists of two porous nickel electrodes separated by a capillary membrane which holds the potassium hydroxide (KOH) electrolyte (Fig. 2). Rigid nickel-plated magnesium electrode holders support the electrodes, manifold the reactant gases to the electrodes, and conduct the electrical energy from the cell. Reactants are consumed only when a load is connected to the cell.

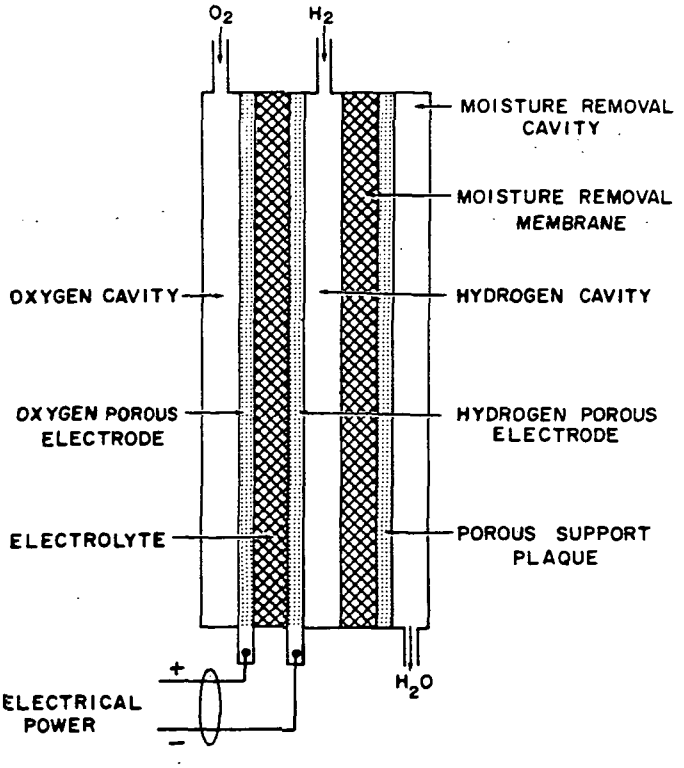


FIGURE 2. H_2 - O_2 CAPILLARY FUEL CELL

The most important feature of this cell is the capillary membrane, a sheet of very pure asbestos. The function of this membrane is to hold the electrolyte in contact with the electrodes. The capillary potential of the membrane (the differential pressure required to force electrolyte from the largest pore) is many times greater than the normal operating pressure of the cell. This property of the capillary membrane makes the cell independent of gravity, produces a uniform separation of electrodes and of the reactant gases, and allows communication of water and hydroxyl ions between the electrodes. The electrode holders compress the membrane between the electrodes, thus forming a compact cell that is shock, vibration, and acceleration resistant. The membrane is inert and is unaffected by KOH or repeated thermal cycling.

STATIC MOISTURE REMOVAL

As water is produced in the cell, a reduction of the electrolyte concentration results. The highest voltage is obtained when the electrolyte concentration is between 33 percent and 38 percent KOH. The cells are generally built with an initial concentration of approximately 35 percent KOH. The concentration can be varied in relation to the volume of electrolyte used to obtain the desired conditions.

For a given concentration of KOH, the water vapor pressure in the adjacent reactant gas cavities will be related to the temperature of the electrolyte. For a given temperature, the vapor pressure increases with a reduction of KOH concentration. The water vapor molecules seek equilibrium with the reactant gas molecules in the gas cavities, and the sum of the partial pressures of the gases in a cavity equals the total operating gas pressure of the cell.

The increase in vapor pressure at constant temperature, with a reduction of electrolyte concentration (caused by an increase in the amount of water produced in the cell), provides the basis for the controlled removal of the product water. Water removal is accomplished by placing a second asbestos membrane (the water transport membrane), saturated with a KOH solution of slightly higher concentration than the electrolyte holder, directly in contact with one of the reactant gas cavities as shown in Figure 2.

The KOH in this membrane absorbs some of the water vapor in the gas cavity. This asbestos membrane is backed up with a porous nickel support plaque (similar in construction to an electrode), which serves

as a reservoir for the solution volume change associated with a variation of KOH concentration in the water removal membrane. Since the water transport membrane is saturated with the KOH solution, the loss of reactant gas by diffusion into the asbestos is negligible.

The water absorbed in the water transport membrane reduces the KOH concentration and increases the volume. The cavity adjacent to the support plaques is filled only with water vapor. The vapor pressure is related to the KOH concentration in the membrane and the cell temperature. By lowering the cavity pressure to the desired vapor pressure, water is withdrawn from the cell as vapor. Control of KOH concentration in the two membranes and of the water removal rate from the cell at a given temperature is thus directly dependent on the water removal cavity pressure.

This system is ideal for space applications. The water vapor can be exhausted directly to space vacuum; the only hardware required for the electrolyte moisture control system is a control valve set at a relief pressure associated with the operating temperature. With the addition of some simple auxiliary equipment, the water can be recovered for human consumption. It is possible to operate the cells within a range of temperatures by adjusting the cavity pressure control valve. This adjustment would be automatic with change in temperature and would result in a very stable water removal system throughout the selected temperature range.

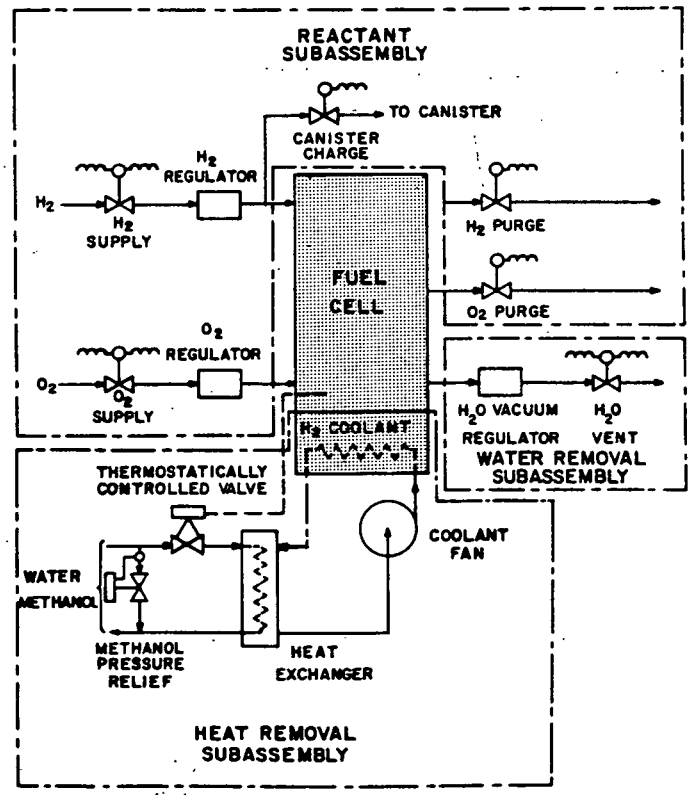


FIGURE 3. REACTANT AND FLUID FLOW

The reactant gas cavities are normally "dead ended" and require only an occasional purge (dependent on the purity of the gas supplied) to remove inert substances. This system is completely static with the only moving components being accessories such as solenoid and regulating valves as shown in Figure 3. The system lends itself to low gravity conditions since fluid transport occurs in the vapor state, and there are no free liquids or gravity-dependent components.

POTABLE WATER RECOVERY

The potable water produced in the fuel cell and removed by the static moisture removal system can be recovered for use, if desired. Recovery of the water would, however, increase the problem of heat removal since the waste heat normally dissipated by the ejection of the water vapor in space would have to be removed by other means.

Reclamation of the water would add mass and volume to the fuel cell assembly, e.g., equipment to condense and store the water and larger radiators to dissipate waste heat.

There are several methods feasible for water recovery. One method that has operated successfully utilizes a capillary membrane condenser.

The condenser is connected to the vacuum control valve in the moisture removal system and is evacuated to a lower pressure than the desired water removal cavity pressure. A suitable coolant supply is required to maintain the condenser temperature below the temperature at which the vapor pressure of pure water is equal to the water removal cavity vapor pressure.

As water vapor is released by the vacuum control valve, the vapor is condensed on a capillary asbestos membrane and transported through the membrane by its capillary action. The water collects in a cavity on the opposite side of the membrane and is transferred to storage by a small pump. Condensation of the vapor on the membrane maintains the evacuated pressure in the condenser.

FUEL CELL POWER SUPPLY

For integration into a space vehicle, a fuel cell power supply must be a complete, self-sustained system, capable of furnishing the power demands of the electrical distribution system during prelaunch and space operations. A survey of various NASA mission studies has shown that a typical electrical system consists of two or more electrical 28-volt buses, each requiring 2 kW or less. On this basis, the selection of a fuel cell assembly with a nominal 2 kW rating is appropriate. Therefore, a typical fuel cell power system will consist of two or more fuel cell assemblies, reactant storage and distribution assemblies, a coolant system, and the interfacing electrical and mechanical systems.

Each fuel cell assembly will consist of a canisterized fuel cell module, auxiliary assemblies, and subassemblies to provide the required electrical power. The fuel cell assembly with the various electromechanical subsystems, as described, will be self-contained within a cylindrical canister approximately 30 cm (11.5 in.) in diameter and 75 cm (30 in.) in length (Fig. 4). The assembly will have a mass of 41 kg (90 lbs) and a maximum continuous rating of 2200 watts. It can deliver electrical power from 830 to 2600 watts at a self-regulated voltage of 28 ± 2 volts.

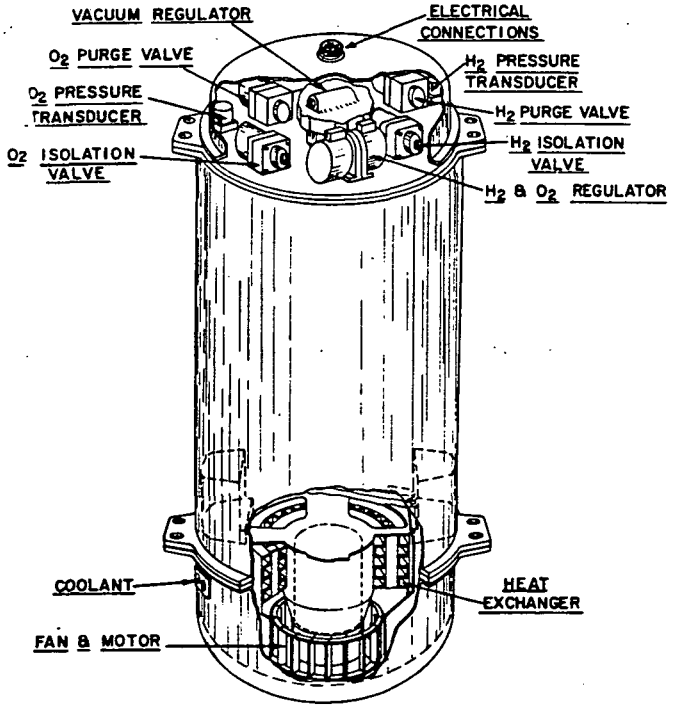


FIGURE 4. FUEL CELL ASSEMBLY

Each assembly will have a stack, or battery, of 66 individual cells (two parallel sets of 33 series-connected cells as shown in Figure 5). Each cell consists of the following:

a. Electrodes — porous rectangular plaques of sintered nickel, 0.061 cm (0.024 in.) thick and plated with platinum and palladium catalysts.

b. Electrolyte-holding membrane — a sheet of very pure grade of asbestos, 0.08 cm (0.030 in.) thick, saturated with a solution of 35 percent KOH electrolyte.

c. Moisture removal membrane — identical to electrolyte-holding membrane except initial saturation of KOH is 40 percent.

d. Support plaques — identical to electrolyte-holding membrane except support plaques have no electrical connection to the bus system.

The reactant control subsystem and the water removal subsystem are mounted on one end plate of the fuel cell reactor subsystem (Fig. 4). The reactant control subsystem consists of inlet valves for each gas, a pressure-balancing regulator, purge rate control, and purge valve for each reactant. The water removal subsystem consists of a vacuum regulator to maintain desired water vapor pressure and a solenoid valve to isolate the vacuum regulator from the atmosphere when the fuel cell is not operating.

Also contained in the fuel cell assembly are the heat removal and the heater element subsystems. Heat removal is accomplished by circulating cooled hydrogen around fins in the fuel cell reactor subsystem. The heater element subsystem is used to preheat the fuel cell assembly and to maintain operating temperatures under low load conditions.

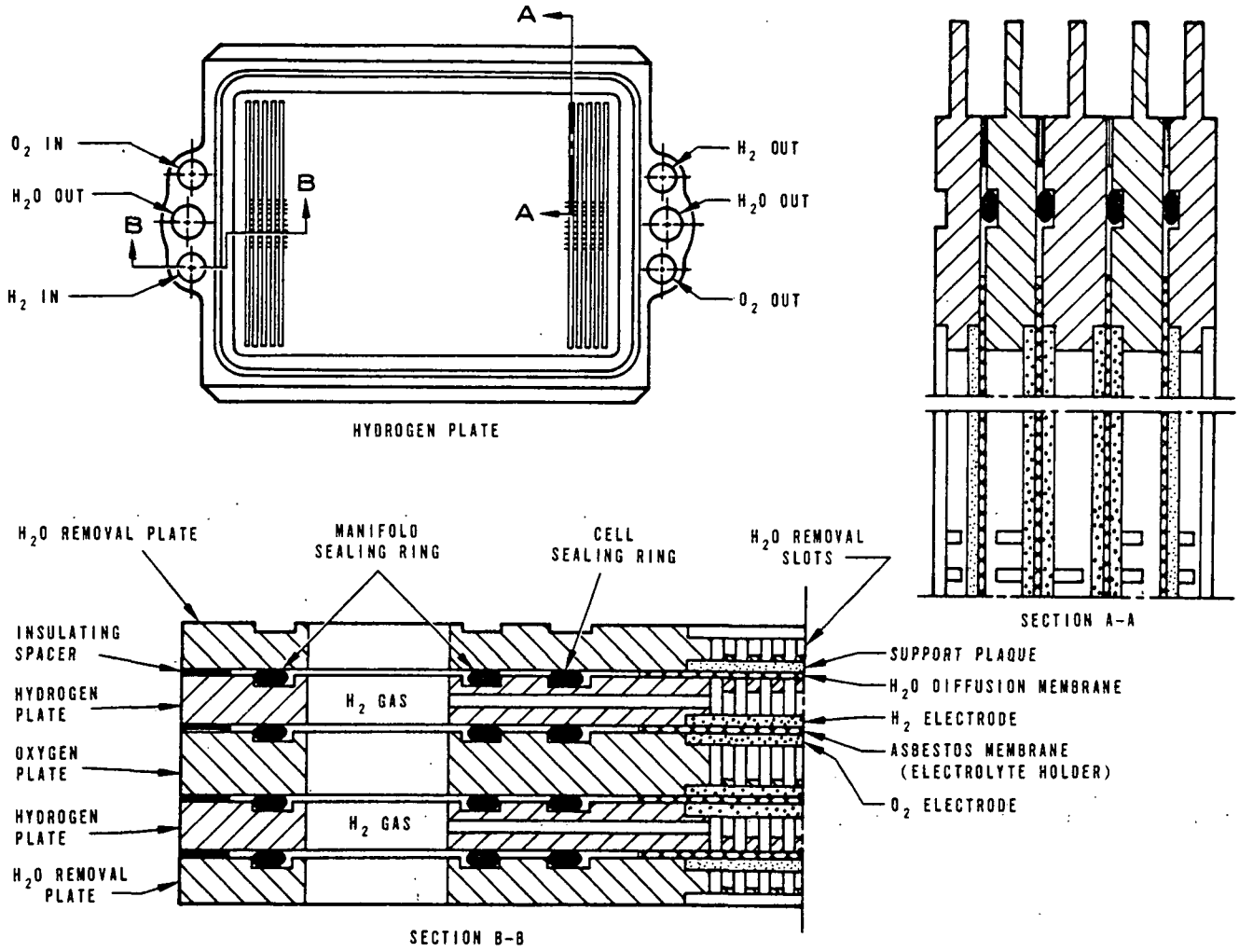


FIGURE 5. CROSS SECTION OF CELL MANIFOLDING

ELECTRICAL CONTROL

The design of the electrical control system is largely dependent on the actual application and vehicle design. Typically, it incorporates the following functions:

a. Power switching
b. Overcurrent
c. Voltage regulation
d. Purging
   (1) Scheduled
   (2) Non-scheduled
e. Temperature control
f. Reactant control
g. Manual remote control (for testing and prelaunch operations)

Power isolation and overcurrent protection are standard requirements for electrical power systems for space vehicles.

The fuel cell has inherently good voltage regulation under load by virture of its flat volt-ampere curve. Where closer voltage regulation is required, other techniques may be employed such as direct dc to dc conversion. The type of voltage regulation used will depend on the specific application and ultimate use of the system power output.

The scheduled-purge control senses the canister output (ampere-hours) and purges the oxygen and hydrogen cavities of inert substances after a predetermined output has been accumulated. The non-scheduled purge control senses the cell voltage and permits additional purging when a drop in voltage indicates the need for additional removal of inert substances.

Normal temperatures are maintained by control of the heater and by a proportional valve in the coolant system. Below-normal temperatures cause a thermostatic switch to connect the heater to the fuel cell bus. The heater would be required only for startup or for very low power output operation during prelaunch. The thermal design of the fuel cell assembly will allow standby operation in space with only 60 to 80 watts of parasitic power required to maintain temperature.

The reactant pressure is automatically controlled by the pressure regulator. Solenoid valves are used, however, to isolate the canisters from the reactant supplies and to open and close the vent lines both for automatic and, if desired, manual operation.

Manual override of automatic functions can be provided for a manned space vehicle application.

REACTANT SUPPLY

To minimize the volume (and therefore the container mass) of the reactant gas supplies, hydrogen and oxygen should be stored in the supercritical stage. For long missions, storage of the reactants as solid blocks may produce further mass and volume economies and should be considered.

Inert substances, from the reactant supply, are not harmful to the fuel cell chemically but do affect its operation by accumulating in the pores of the electrode. Argon, normally found in small quantities in oxygen produced by air liquefaction and distillation processes, is typical of these substances. As oxygen is consumed through the pores of the cathode, the argon concentration increases because it does not enter into the reaction and accumulates in the electrode pores. As the argon concentration increases, the oxygen concentration, of necessity, decreases. Periodic purging of the cells removes the accumulated inert substances and prevents deterioration of the cell.

Contaminants that are considered harmful are acidic anions which might be ionized in the cell. These include the halogens, sulfur, nitrogenous oxides, and carbonaceous oxides. The contaminants are not usually found in analytical grade reagents.

For any reactant supply purity, a purge interval exists for which the total system mass is a minimum. This purge is more conveniently expressed as a time-averaged purge rate. For a reactant purity of 0.999, the optimum purge rate is about 3 percent of the stoichiometric consumption.

Reactant gas purge duration and frequency are controlled by the electrical control system. Normally-closed solenoid valves in each reactant outlet of the fuel cell assembly are opened and closed as required by the purge controller.

Manual purge can also be provided to allow override of the purge controller for manned vehicles.

THERMAL CONTROL

Thermal control is required for efficient operation of the fuel cell and the static moisture removal system. Waste heat is produced as a by-product of the reaction in the fuel cell at the electrodes and in the electrolyte. Some waste heat also results from parasitic power consumption by fuel cell auxiliary components such as relays and solenoid valves.

The fuel cell is designed to operate within a range of 85 to 100° C. This temperature range provides a specific power output high enough to keep mass relatively low while avoiding high-temperature mechanical and material problems.

CONCLUSIONS

The results of the research and development effort have been most gratifying. The program has progressed to the point where design and development necessary to incorporate the equipment into the vehicle system are required.

The concept of vapor pressure control is considered a significant advancement since it minimizes the moisture removal and control problems. The system is inherently stable and suitable for space applications. Where water reclamation is not required, the static moisture removal system simplifies the heat removal problem by removing a large portion of the heat when the water vapor is ejected to space.

The equipment designs are considered rugged and reliable, simple to manufacture and assemble, and readily adaptable to space vehicles.

III RESEARCH AND SUPPORTING ENGINEERING...

A. STATIC LOAD AND FLOW EQUATIONS FOR A DOUBLE-ACTING, JOURNAL-FED, EXTERNALLY-PRESSURIZED THRUST BEARING by Peter H. Broussard, Jr., John Burch, and Emmett Martz

The load and flow equations for a double-acting, journal-fed, externally pressurized thrust bearing are derived from the Navier-Stokes equations. Load displacement and flow curves are then given as functions of the bearing parameters. The theory is compared with experiments; the agreement is excellent and provides another instance of the effectiveness of the Navier-Stokes equations in describing gas film lubrication.

INTRODUCTION

The type of externally-pressurized gas bearing used to provide a near-frictionless support for a gyroscope gimbal axis is analyzed. Within its capacity, the double-acting bearing supports applied loads in any direction; however, only loads applied in the axial direction are treated here.

For brevity, the complete bearing, consisting of a double-acting thrust bearing and a 360-degree journal bearing, will be referred to as simply a "bearing."

Generally in assemblies of this type, the journal and thrust bearings are fed separately through restrictors and have a common exhaust. Consequently, the bearings can be analyzed separately since, in practice, the influence of one bearing on the other is negligible.

The journal-fed bearing retains the readily identified thrust and journal bearings but differs in an important respect from the separately fed bearing because gas is fed only to the journal, as the name implies. The thrust bearing utilizes the gas flowing from the journal bearing; that is, the journal is made to exhaust through the thrust bearing.

In a double-acting, journal-fed bearing, a pressure gradient is maintained across the thrust bearing by the action of viscous shear forces retarding the flow of gas in the gaps formed by the surfaces of the bearings. With the cylinder in an eccentric position, some gas from the lower row of restrictors must flow through the bottom thrust bearing. This flow, in turn, establishes a pressure gradient and the exact relationships are derived. As one would expect, the length of the journal and the journal clearance are sensitive factors in the performance of the thrust bearing.

The journal-fed bearing has several advantages over the conventional bearing:

1. The gas consumption is only about 50 percent as much as a bearing which has conventional feeding end plates. Although the ultimate axial and radial load carrying capacity of the journal-fed bearing is less than that of the conventional bearing, the load carrying capacity per unit volume of gas is much higher for the journal-fed bearing than the conventional bearing.

2. The turbine torque caused by the end plates in a journal-fed bearing is less than that caused by the end plates in a conventional bearing. This reduction of turbine torque contributes towards making the guidance system more accurate.

3. The journal-fed bearing has a greater axial dynamic load carrying capacity than a conventional bearing.

4. The dynamic stability of the journal-fed bearing is as good as or better than that of a conventional bearing.

5. The end plates of a journal-fed bearing have no gas feeding system and therefore are much simpler and cheaper to manufacture than the end plates of a conventional bearing.

The solution to the problem of calculating static load carrying capability and flow is found in the use of the reduced Navier-Stokes equations as applied to a gas flowing in a slot. Since only the static case is being considered, there is no need to use the Reynolds

equation. The bearing surfaces possess neither normal nor lateral relative motion; therefore, neither the hydrodynamic nor squeeze forces must be considered. In this analysis, the reduced Navier-Stokes equations are solved without feeding, and then the feeding is introduced through the boundary conditions.

ANALYSIS

The analysis is concerned with the flow of a gas in a narrow gap. The gaps under consideration are those formed by the surfaces of the float and the surfaces of the sleeve and end plate. The dynamic behavior of a gas is described by the well-known Navier-Stokes equations. However, for the configuration considered here, certain realistic assumptions can be made which, when applied, reduce the Navier-Stokes equations to a more tractable form.

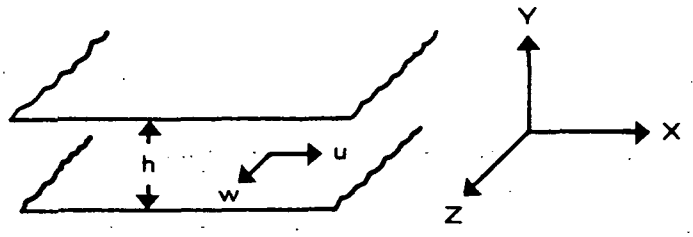


FIGURE 1. SKETCH OF GAP

Consider the gap shown schematically in Figure 1 with a Cartesian coordinate system. The surfaces of the gap are assumed to have neither lateral nor normal motion relative to each other; u is the velocity of the gas in the x-direction and w is the velocity of the gas in the z-direction. The velocity in the y-direction is assumed to be zero.

The following assumptions are made:

1. The height, h, of the gap is very small compared with the length and breadth of the gap (approximately 10^{-3} compared to unity in this case).

2. The pressure, p, is constant across the film.

3. The flow is laminar.

4. No external forces act on the film.

5. Inertia forces are small compared to viscous shear forces.

6. The no-slip condition holds at the bearing surfaces.

7. All other velocity gradients are negligible compared with $\frac{\partial u}{\partial y}$ and $\frac{\partial w}{\partial y}$.

If these conditions are applied to the Navier-Stokes equations, the two following equations result. All resulting equations are derived directly from these fundamental equations (A, A', B, and B') in conjunction with an equation giving the flow through a restrictor. (See Ref. 1 for a complete treatment of the Navier-Stokes equation and its reduction to the Reynolds lubrication equation. A derivation of the reduced Navier-Stokes equations from first principles is given in the Appendix).

$$\frac{\partial^2 u}{\partial y^2} = \frac{1}{\mu}\frac{\partial p}{\partial x} \qquad (A)$$

$$\frac{\partial^2 w}{\partial y^2} = \frac{1}{\mu}\frac{\partial p}{\partial z}. \qquad (B)$$

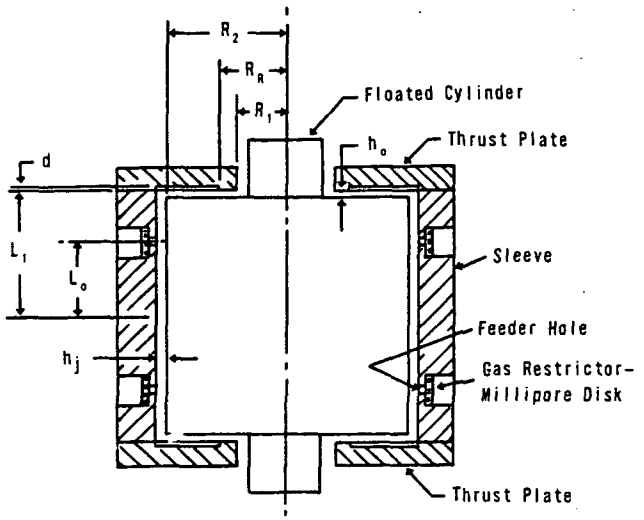
In cylindrical coordinates, these equations become:

$$\frac{\partial^2 v_r}{\partial y^2} = \frac{1}{\mu}\frac{\partial p}{\partial r} \qquad (A')$$

$$\frac{\partial^2 v_\theta}{\partial y^2} = \frac{1}{\mu r}\frac{\partial p}{\partial \theta} \qquad (B')$$

where v_r and v_θ are the velocities in the r and θ directions, respectively, and μ is the viscosity of the gas. Only equations A and A' will be used in this discussion. The flow in the journal bearing will be considered first.

In the journal, gas flows through a restrictor (in this case, millipore) into the journal gap (Fig. 2). Since the float is in a concentric radial position, only small pressure gradients are present circumferentially, so the flow is predominately in an axial direction. A circumferential pressure gradient exists in the immediate neighborhood of the feeder hole, but it disappears a short distance away. Neglect of this gradient seems to introduce little error especially since the thrust plate is being considered. Thus, the gas will be considered as flowing axially down the journal gap. In particular, if n feeder holes are in a feeder plane, it will be assumed that the gas from one feeder hole flows in a channel of height h_j and width, $\frac{2\pi R_2}{n}$, where R_2 is the radius of the float. The slight curvature of the channel can be neglected. Use is made of equation A of the Navier-Stokes equations to



R_1 = inner radius of thrust plate
R_R = inner radius of recessed annulus
d = depth of recessed annulus
R_2 = radius of floated cylinder
L_0 = the distance from the midpoint of the sleeve to either row of feeder holes
L_1 = the distance from the midpoint of the sleeve to either end
h_j = the radial clearance = the sleeve radius minus R_2
h_0 = the axial clearance = L_1 minus half the length of the floated cylinder

FIGURE 2. CROSS SECTION OF BEARING

calculate the flow and pressure distribution in a rectangular gap.

After flowing down the journal gap from the feeder hole to the junction of the journal and thrust bearing, the gas flows to exhaust in a convergent manner. The gas enters the thrust gap (in an extremely uniform manner) when the float is in concentric journal position. Pressure measurements around the circumference of the journal, immediately prior to its junction with the end plate, show no pressure gradient, Thus, the prevailing conditions at the end plate gap entrance are exactly those assumed in the idea of line feeding; i.e., the gas emanates uniformly from a line, streamlines have no curvature, and pressure gradients normal to streamlines are zero.

A. FLOW IN A SLOT

A cross section of the slot is shown in Figure 3. It has a height, h, and width, w; both possibly varying with x. Use is made of equation A:

$$\frac{\partial^2 u}{\partial y^2} = \frac{1}{\mu}\frac{\partial p}{\partial x}. \qquad (1)$$

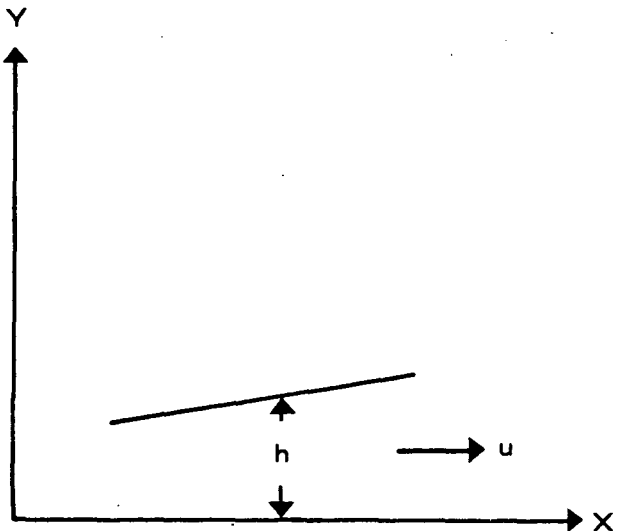


FIGURE 3. CROSS SECTION OF SLOT

The flow is assumed to be in the x-direction. Therefore $\frac{\partial p}{\partial z} = 0$, and equation B will not be used. Also, since $\frac{\partial p}{\partial y} = 0$, then $\frac{\partial p}{\partial x} = \frac{dp}{dx}$.

Since p is not a function of y, equation 1 can be integrated twice with respect to y to give

$$u = \frac{1}{2\mu}\frac{dp}{dx}\, y^2 + c_1 y + c_2. \qquad (2)$$

The bearing surfaces are not moving relative to each other and c_1 and c_2 can be evaluated from

$u = 0$ at $y = 0$
$u = 0$ at $y = h$ } no slip at surface.

This implies

$c_2 = 0$

and

$$c_1 = -\frac{1}{2\mu}\frac{dp}{dx}\, h$$

so that equation 2 becomes

$$u = \frac{1}{2\mu}\left(-\frac{dp}{dx}\right)\left(h\,y - y^2\right). \qquad (3)$$

Thus, the velocity profile is a parabola in y. The average velocity across any cross section is

$$u_{avg} = \frac{1}{h}\int_0^h u\,dy = -\frac{dp}{dx}\,\frac{h^2}{12\mu}.$$

The volume flow rate, Q, at any cross section of the slot can now be calculated:

$$Q = u_{avg} hw = -\frac{wh^3}{12\mu}\frac{dp}{dx}. \tag{4}$$

In exactly the same way, an analogous flow equation can be derived in cylindrical coordinates:

$$Q = -\frac{wh^3}{12\mu}\frac{dp}{dr} \tag{5}$$

where, as before, w and h may be functions of the downstream distance (in this case, r).

At this point, it will be assumed that the gas obeys the perfect gas law; i.e., it satisfies $p = \rho RT$, where ρ = density, R = universal gas constant, and T = absolute temperature. Also, it will be assumed that the thermodynamic process is isothermal. These assumptions, along with the continuity equation, imply that for any two cross sections 1 and 2 of a channel, if p_1, Q_1 and p_2, Q_2 are the pressure and volume flow rates at 1 and 2, respectively, then they are related by

$$p_1 Q_1 = p_2 Q_2. \tag{6}$$

B. FLOW AND PRESSURE DISTRIBUTION IN A RECTANGULAR SLOT OF CONSTANT WIDTH AND HEIGHT

Figure 4 shows gas flowing in a rectangular slot of length $L_1 - L_0$, width, w, and height, h. Subscript "0" refers to quantities at $x = L_0$ and subscript "1" refers to quantities at $x = L_1$. The gas enters the slot at pressure p_0 and volume flow rate Q_0. At a distance x downstream, the pressure is p and volume flow rate is Q; these are related by

$$pQ = p_0 Q_0$$

or

$$Q = \frac{p_0 Q_0}{p}. \tag{7}$$

Substitution of Q in equation 7 into equation 5 gives

$$\frac{Q_0 p_0}{p} = -\frac{wh^3}{12\mu}\frac{dp}{dx}. \tag{8}$$

Equation 8 can be integrated to

$$\frac{p^2}{2} = -\frac{12\mu Q_0 p_0}{wh^3}\,x + C$$

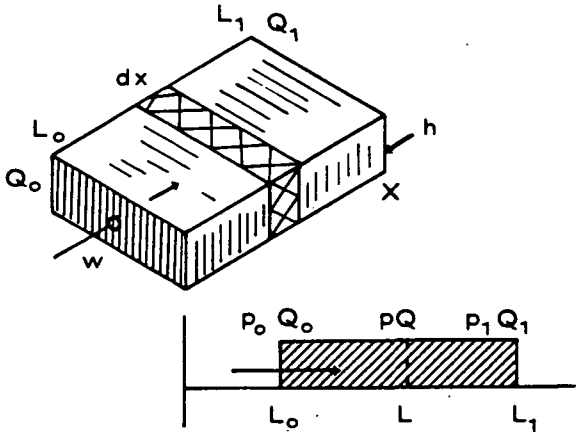


FIGURE 4. FLOW IN RECTANGULAR SLOT

and C can be evaluated from the condition $p = p_0$ at $x = L_0$. If this is done, the following equation results.

$$\frac{p^2}{2} = -\frac{12\mu Q_0 p_0}{wh^3}\,x + \frac{p_0^2}{2} + \frac{12\mu Q_0 p_0 L_0}{wh^3} \tag{9}$$

which gives the pressure distribution in terms of flow, dimensions, viscosity, and initial pressure. The flow rate, Q_0, can be determined from equation 9 and the condition $x = L_1$, $p = p_1$. If this condition is applied to equation 9, Q_0 can be calculated

$$Q_0 = \frac{wh^3\,(p_1^2 - p_0^2)}{24\mu p_0\,(L_0 - L_1)}. \tag{10}$$

Q_0, then, is the volume flow rate at pressure p_0.

If Q_0 in equation 10 is used in equation 9, the following expression for the pressure distribution results.

$$p^2 = p_0^2 + (p_1^2 - p_0^2)\,\frac{L_0 - x}{L_0 - L_1}$$

which can be written as

$$p = p_0 \left\{1 + \frac{(L_0 - x)}{(L_0 - L_1)}\left[\left(\frac{p_1}{p_0}\right)^2 - 1\right]\right\}^{\frac{1}{2}}. \tag{11}$$

Thus, the pressure distribution is determined once the length of the channel and the entrance and exit pressures are known.

C. FLOW AND PRESSURE DISTRIBUTION IN AN ANNULUS OF CONSTANT HEIGHT

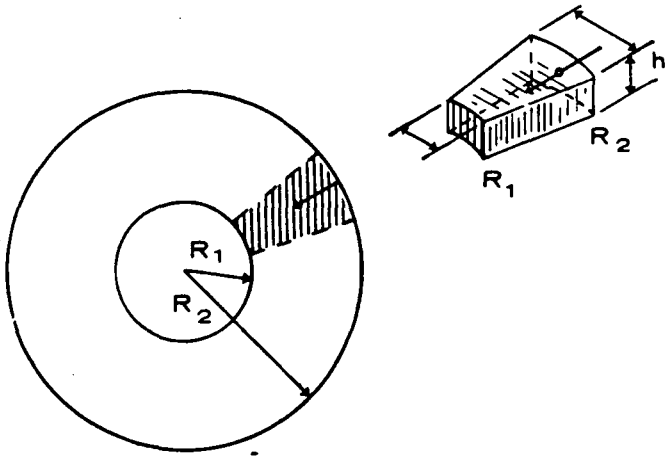


FIGURE 5. FLOW IN ANNULUS

Figure 5 shows a sketch of the annular thrust bearing. Gas enters the annular gap uniformly at the outer periphery ($r = R_2$) at pressure p_0 and flows radially inward and exhausts at $r = R_1$ and pressure p_1. Equation A' is used and the treatment is the same as that for the rectangular channel except that the width, w, is not a constant, but rather the "width" is $w = 2\pi r$. With this change, the equation can be integrated with the following results for flow rate and pressure distribution.

$$Q_0 = \frac{-\pi h^3}{12\mu \ln \frac{R_1}{R_2}} \frac{p_1^2 - p_0^2}{p_0} \quad \text{(flow rate at pressure } p_0\text{)} \tag{12}$$

$$p = p_0 \left\{ 1 + \frac{\ln \frac{r}{R_2}}{\ln \frac{R_1}{R_2}} \left[\left(\frac{p_1}{p_0} \right)^2 - 1 \right] \right\}^{\frac{1}{2}} . \tag{13}$$

Note that the pressure distribution meets the boundary conditions: at $r = R_1$, $p = p_1$ and at $r = R_2$, $p = p_0$.

D. MILLIPORE RESTRICTOR

The gas generally enters the bearing through some type of restrictor. Orifices, porous metal plugs, capillary tubes, and millipore paper disks are commonly used. Only millipore is considered here.

An analytical expression relating the flow rate with the supply pressure, p_S, the exhaust pressure, p_0, and the type of millipore is necessary to develop an analytical model of the gas bearing. No theoretical equation, based on fundamental quantities (porosity, area, etc.), is known to exist. However, excellent agreement can be obtained by assuming a relation of the form

$$Q = k \frac{p_{ref}}{p_0} (p_s^2 - p_0^2). \tag{14}$$

The constant k is obtained by choosing two points on the measured millipore flow-pressure curve and by referring to some arbitrary reference pressure. p_{ref} (normally atmospheric pressure) is the pressure at which the volume flow rate Q is measured. The constant, k, determined in this manner, will hold only for the type of millipore tested.

E. APPLICATION OF PRECEDING PRESSURE AND FLOW EQUATIONS TO A DOUBLE-ACTING, JOURNAL-FED THRUST BEARING

A cross section giving the pertinent dimensions of a double-acting, journal-fed thrust bearing is shown

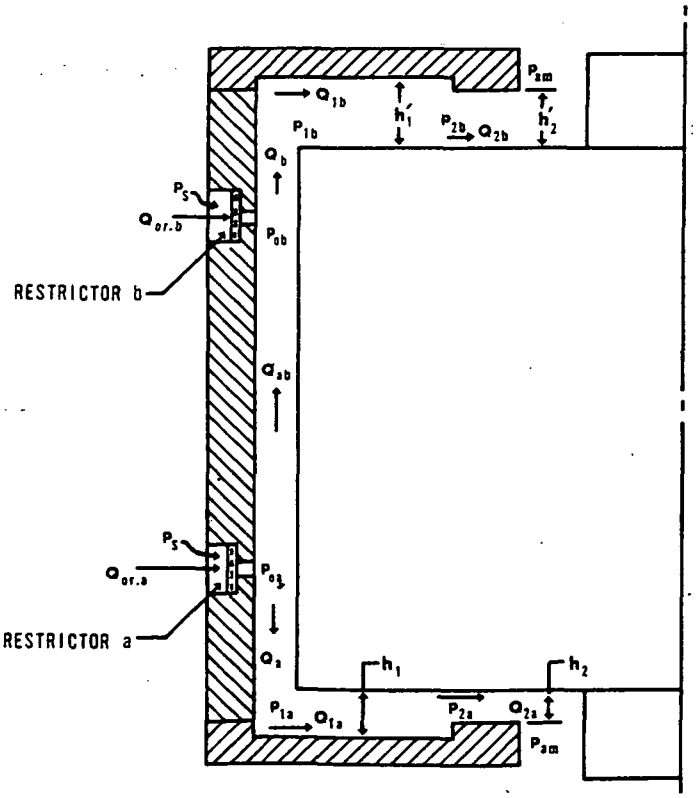


FIGURE 6. CROSS SECTION OF PORTION OF BEARING SHOWING FLOW, PRESSURE, AND DIMENSIONS

in Figure 6. The bearing is unique because no restrictors are located in the thrust plates themselves; flow from the journal is utilized for feeding the thrust plates. While the journal bearing is not treated here, it is capable of supporting a load in the radial direction also. The purpose in using this arrangement is the great reduction in gas consumption, an important consideration in long space missions.

The bearing shown in Figure 6 is assumed to be loaded on the lower side. As indicated, gas enters the bearing through the restrictors a and b; the flow rates through the restrictors are designated by $Q_{or\,a}$ and $Q_{or\,b}$, respectively. For the bearing to develop a load carrying capacity, a greater pressure drop (to ambient pressure) must be across the bottom end plate than across the top end plate. This condition in a journal-fed thrust bearing is made plausible because, in an eccentric float position, the flow through the bottom gap is greatly restricted and the flow through the top gap is greatly relieved (equation 12 relates the flow with pressure and gap height). Thus, there will be a pressure increase at the bottom thrust bearing and a pressure decrease at the top thrust bearing.

To develop a mathematical expression for the static load carrying capacity and flow for the bearing will require the equations previously developed. Figure 6 is a sketch of a portion of the bearing, showing the relevant pressures, flow rates, and gaps.

These quantities are defined as follows:

W = applied axial load

h_2 = gap height between the floated cylinder and the loaded thrust plate from $r = R_1$ to R_R

h_1 = gap height between the floated cylinder and the loaded thrust plate from $r = R_R$ to R_2; therefore $h_1 = h_2 + d$

h_2' = gap height between the floated cylinder and the unloaded thrust plate from $r = R_1$ to R_R (note: $h_2 + h_2' = 2h_0$, h_0 = clearance)

h_1' = gap height between the floated cylinder and the unloaded thrust plate from $r = R_R$ to R_2; therefore $h_1' = h_2' + d$

p_s = supply pressure

p_{am} = ambient pressure

p_{oa} = pressure at the feeder hole entry into the radial gap for the row of feeder holes nearest the loaded thrust plate

p_{ob} = pressure at the feeder hole entry into the radial gap for the row of feeder holes nearest the unloaded thrust plate

p_{1a} = pressure in the loaded axial gap at $r = R_2$

p_{1b} = pressure in the unloaded axial gap at $r = R_2$

p_{2a} = pressure in the loaded axial gap at $r = R_R$

p_{2b} = pressure in the unloaded axial gap at $r = R_R$

$Q_{or\,a}$ = flow through restrictor a

$Q_{or\,b}$ = flow through restrictor b

Q_{ab} = flow through gap from restrictor a to restrictor b

Q_a = flow through gap from restrictor a to junction of journal and bottom end plate

Q_b = flow through gap from restrictor b to junction of journal and top end plate

Q_{1a} = flow from junction to step in bottom end plate

Q_{1b} = flow from junction to step in top end plate

Q_{2a} = flow from step in bottom end plate to ambient pressure

Q_{2b} = flow from step in top end plate to ambient pressure

For volume flow rate to be used correctly, the pressure at which the flow rate is measured must be specified. This is clear from the context of the formulas. Furthermore, flow at one pressure can be converted to flow at another pressure simply by equation 6, which is used in the following problem, without specific comment necessarily being made.

The solution of the problem lies in the continuity of flow through the bearing; i.e., the flow from the restrictors must be accounted for. Specifically, flows are equated in adjoining gaps, thus allowing a relationship to be developed between quantities (e.g., pressures and dimensions) in one gap to similar quantities

in the other gap. As an example, the flow, Q_a, down the journal gap is equated to the flow, Q_{1a}, in the recessed convergent gap in the thrust gap; both flows are computed at the same pressure.

Using the dimensions and notation of Figures 2 and 6, the flow equations are a direct consequence of equations 10, 12, and 14:

$$Q_{or\ a} = k \frac{p_{ref}}{p_{0a}} (p_s^2 - p_{0a}^2) \tag{15}$$

$$Q_{or\ b} = k \frac{p_{ref}}{p_{0b}} (p_s^2 - p_{0b}^2) \tag{16}$$

$$Q_{ab} = \frac{B}{A} \frac{(p_{0b}^2 - p_{0a}^2)}{p_{0a}} \tag{17}$$

$$Q_a = \frac{B}{D} \frac{(p_{1a}^2 - p_{0a}^2)}{p_{0a}} \tag{18}$$

$$Q_b = \frac{B}{D} \frac{(p_{1b}^2 - p_{0b}^2)}{p_{0b}} \tag{19}$$

$$Q_{1a} = E \frac{(p_{2a}^2 - p_{1a}^2)}{p_{1a}} \tag{20}$$

$$Q_{2a} = F \frac{(p_{am}^2 - p_{2a}^2)}{p_{2a}} \tag{21}$$

$$Q_{1b} = \frac{E' (p_{2b}^2 - p_{1b}^2)}{p_{1b}} \tag{22}$$

$$Q_{2b} = \frac{F' (p_{am}^2 - p_{2b})}{p_{2b}} \tag{23}$$

where

$$A = 48\mu L_0$$

$$B = -2\pi R_2 h_j^3$$

$$D = 24\mu (L_1 - L_0)$$

$$E = \frac{\pi h_1^3}{12\mu \ln \frac{R_R}{R_2}} = \frac{\pi (h_2 + d)^3}{12\mu \ln \frac{R_R}{R_2}}$$

$$F = \frac{\pi h_2^3}{12\mu \ln \frac{R_1}{R_R}}$$

$$E' = \frac{\pi h_1'^3}{12\mu \ln \frac{R_R}{R_2}} = \frac{\pi (2h_0 - h_2 + d)^3}{12\mu \ln \frac{R_R}{R_2}}$$

$$F' = \frac{\pi h_2'^3}{12\mu \ln \frac{R_1}{R_R}} = \frac{\pi (2h_0 - h_2)^3}{12\mu \ln \frac{R_1}{R_R}}.$$

From equation 6, the following equations can be written (where the flow rates are measured at the appropriate pressures).

$$p_{1a} Q_{1a} = p_{2a} Q_{2a} \tag{24}$$

$$p_{1a} Q_{1a} = p_{0a} Q_a \tag{25}$$

$$p_{1b} Q_{1b} = p_{2b} Q_{2b} \tag{26}$$

$$p_{1b} Q_{1b} = p_{0b} Q_b \tag{27}$$

$$p_{0a} Q_{or\ a} = p_{0a} (Q_a + Q_{ab}) \tag{28}$$

$$p_{0b} Q_b = p_{0b} (Q_{or\ b} + Q_{ab}). \tag{29}$$

To determine the net load on the thrust plates, p_{1a}, p_{2a}, p_{1b}, and p_{2n} must be known. Equations 24 through 29 are six equations involving the six unknowns p_{1a}, p_{2a}, p_{1b}, p_{2b}, p_{0a}, and p_{0b}. Therefore, expressions for the desired pressure can be found in terms of dimensions, supply, and ambient pressure. The algebra is omitted; only the final solutions are given below.

$$p_{0b}^2 = \frac{U}{V} \tag{30}$$

$$p_{0a}^2 = \frac{BDZp_{0b}^2 - ADkZp_{ref} p_s^2 + ABDEF\ p_{am}^2}{Y} \tag{31}$$

$$p_{1a}^2 = \frac{DEFp_{am}^2 + B(E+F)p_{0A}^2}{Z} \quad (32)$$

$$p_{2a}^2 = \frac{Ep_{1a}^2 + Fp_{am}^2}{E+F} \quad (33)$$

$$p_{1b}^2 = \frac{DE'F'p_{am}^2 + BU(E'+F')p_{0b}^2}{Z'} \quad (34)$$

$$p_{2b}^2 = \frac{E'p_{1b}^2 + F'p_{am}^2}{E'+F'} \quad (35)$$

where

$$U = -AD^2B\,kZZ'p_{ref}p_s^2 - YZ'ADkp_{ref}p_s^2 + AB^2D^2EFZ'p_{am}^2 + ABDE'F'Yp_{am}^2$$

$$V = YZ'BD - YZ'ADkp_{ref} - Z'B^2D^2Z - AB^2Y(E'+F') + ABYZ'$$

$$Y = BDZ - ADkZp_{ref} - AB^2(E+F) + ABZ$$

$$Z = (ED+B)(E+F) - DE^2$$

$$Z' = (E'D+B)(E'+F') - DE'^2.$$

To determine the desired pressures, the following sequence is calculated: Z', Z, Y, V, U, p_{0b}, p_{0a}, p_{1a}, p_{2a}, p_{1b}, and p_{2b}.

Since the pressures have been determined, the forces exerted on the bottom and top of the cylinder can be calculated using the pressure distribution as given by equation 13. The force exerted is $\int pdA$ where dA is an element of area and the function is integrated between the appropriate limits. In polar coordinates, $dA = 2\pi rdr$ in the $r - \theta$ plane. Applying this to the thrust bearing, the total load carrying capacity, W, is

$$W = 2\pi p_{1a}\int_{R_R}^{R_2}\left\{1 + \frac{\ln\frac{r}{R_2}}{\ln\frac{R_R}{R_2}}\left(\frac{p_{2a}^2}{p_{1a}^2} - 1\right)\right\}^{\frac{1}{2}} rdr \quad (36)$$

$$+ 2\pi p_{2a}\int_{R_1}^{R_R}\left\{1 + \frac{\ln\frac{r}{R_R}}{\ln\frac{R_R}{R_1}}\left(\frac{p_{am}^2}{p_{2a}^2} - 1\right)\right\}^{\frac{1}{2}} rdr$$

$$- 2\pi p_{1b}\int_{R_R}^{R_2}\left\{1 + \frac{\ln\frac{r}{R_2}}{\ln\frac{R_R}{R_2}}\left(\frac{p_{2b}^2}{p_{1b}^2} - 1\right)\right\}^{\frac{1}{2}} rdr$$

$$- 2\pi p_{2b}\int_{R_1}^{R_R}\left\{1 + \frac{\ln\frac{r}{R_R}}{\ln\frac{R_R}{R_1}}\left(\frac{p_{am}^2}{p_{2b}^2} - 1\right)\right\}^{\frac{1}{2}} rdr. \quad *$$

To calculate the flow, knowledge of p_{2a} and p_{2b} is sufficient because equation 12 can be used when the gap entrance pressure and gap exit pressure (p_{am}) are known. Applied to bottom and top gaps, equation 12 becomes, respectively,

$$Q = \frac{E}{p_{am}}(p_{2a}^2 - p_{am}^2) \quad (37)$$

$$Q' = \frac{E'}{p_{am}}(p_{2b}^2 - p_{am}^2) \quad (38)$$

and the total flow rate at ambient pressure, Q_t, is the sum of the two:

$$Q_t = Q + Q'.$$

DISCUSSION

The preceding equations were programed on an electronic computer to compare theory with experiment. One experiment that affords a good basis for comparison is a load versus displacement curve. Experimentally, a bearing is loaded and the corresponding gap is measured. Theoretically, the reverse was done; i.e., a gap was assumed and the corresponding force exerted on the bearing was calculated by the equations given.

* These integrals cannot be integrated in closed form, but by appropriate substitutions can be written as probability integrals, which are tabulated.

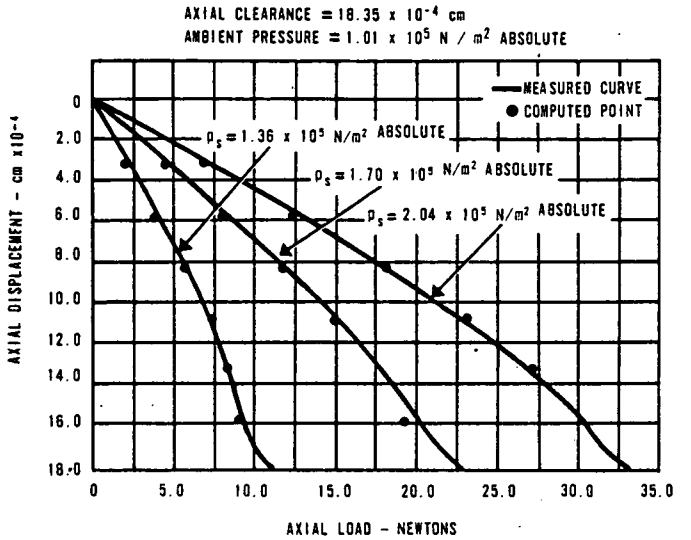


FIGURE 7. AB-3 AXIAL DISPLACEMENT VERSUS LOAD

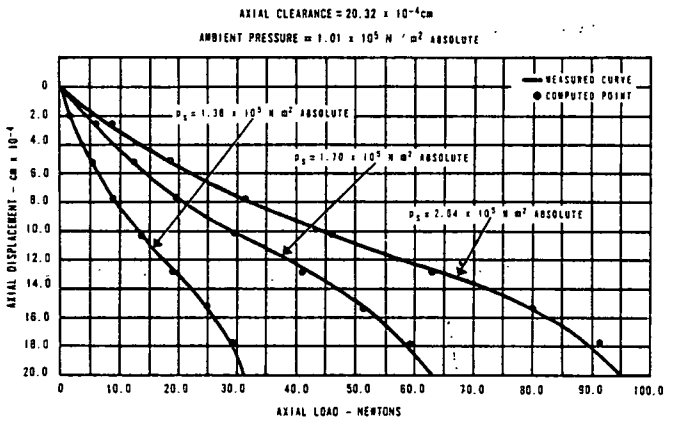


FIGURE 8. AB-5 AXIAL DISPLACEMENT VERSUS LOAD

Figures 7 and 8 show load displacement curves for AB-3 and AB-5 type gas bearings where air was used as the lubricant. The agreement is extremely good for the bearings tested.

Figure 9, a flow rate versus load curve, provides another instance of the validity of the reduced Navier-Stokes equations when applied to the type of bearing under consideration. The curve is self-explanatory and shows that flow rate can also be accurately predicted.

Because of tolerances involved, agreement with experiment will not necessarily be as good as that which is shown here. Naturally, the theoretical equations can be used to study the effect of variation of certain parameters. Some of these parameters are bearing length and diameter, clearances, groove depth and width, position of feeder planes along the journal, supply and ambient pressure, and types of millipore (as characterized by flow rate).

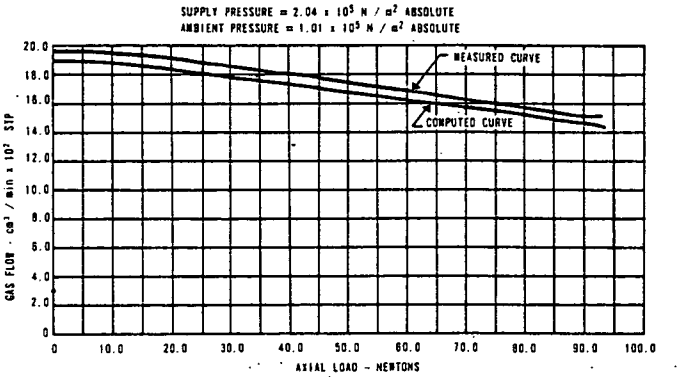


FIGURE 9. AB-5 GAS FLOW VERSUS AXIAL LOAD

While the treatment here applies to the static behavior of the bearing, theoretical load displacement curves can be used to predict the spring constant of the dynamic system. The spring constant, so calculated, can then be used to calculate the undamped resonant frequency of the system. Resonant frequencies calculated in this manner are generally within 5 to 10 percent of their measured value.

Naturally, this procedure is only approximate since a highly nonlinear system is being treated as linear. The approximation is best when the applied excitation is of small amplitude about a concentric position. For large amplitudes about any position or small amplitudes about an eccentric position, it will be necessary to take into account the force generated by the squeeze film.

CONCLUSIONS

The application of the reduced Navier-Stokes equations to the flow of a gas in an externally-pressurized gas bearing, to predict load carrying capacity and flow rate, is amply justified by close agreement between theory and experiment.

APPENDIX

Derivation of Reduced Navier-Stokes Equations

The reduced Navier-Stokes equations are derived by using the assumptions previously made in the analysis.

Referring to the element of fluid in Figure 10, with external forces and inertia neglected and with no pressure gradient in the y-direction, all shear forces τ and pressure forces p in the x-direction are as shown.

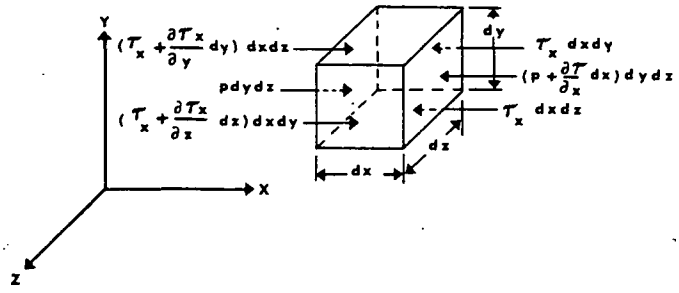


FIGURE 10. PRESSURE AND SHEAR FORCES ACTING ON ELEMENT OF FLUID IN THE X-DIRECTION

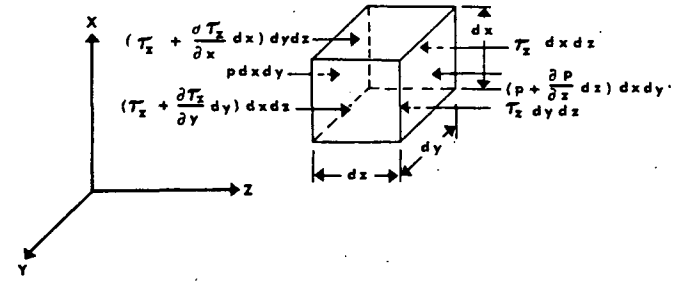


FIGURE 11. PRESSURE AND SHEAR FORCES ACTING ON ELEMENT OF FLUID IN THE Z-DIRECTION

Referring to the same element of fluid in Figure 11, which has the axes rotated, and making the same assumptions, all shear and pressure forces in the z-direction are as shown.

By summing the forces in the x and z direction,

$$pdydz + (\tau_x + \frac{\partial \tau_x}{\partial y} dy)\ dxdz + (\tau_x + \frac{\partial \tau_x}{\partial z} dz)\ dxdy - \tau_x dxdy - \tau_x dxdz - (p + \frac{\partial p}{\partial x} dx)\ dydz = 0 \qquad (1)$$

$$pdxdy + (\tau_z + \frac{\partial \tau_z}{\partial x} dx)\ dydz + (\tau_z + \frac{\partial \tau_z}{\partial y} dy)\ dxdz - \tau_z dxdz - \tau_z dydz - (p + \frac{\partial p}{\partial z} dz)\ dxdy = 0. \qquad (2)$$

By canceling like terms and simplifying,

$$\frac{\partial \tau_x}{\partial y} + \frac{\partial \tau_x}{\partial z} = \frac{\partial p}{\partial x} \qquad (3)$$

$$\frac{\partial \tau_z}{\partial y} + \frac{\partial \tau_z}{\partial x} = \frac{\partial p}{\partial z}. \qquad (4)$$

We now introduce the assumption of a Newtonian lubricant, i.e., a fluid in which the shear stress is proportional to the rate of shear, or

$$\tau_x = \mu \frac{\partial u}{\partial y} \qquad (5)$$

$$\tau_z = \mu \frac{\partial w}{\partial y} \qquad (6)$$

where u = the fluid velocity in the x-direction
w = the fluid velocity in the z-direction
μ = the coefficient of viscosity of the fluid.

By using equations 5 and 6 in equations 3 and 4, respectively,

$$\frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial y \partial z} = \frac{1}{\mu} \frac{\partial p}{\partial x} \qquad (7)$$

$$\frac{\partial^2 w}{\partial y^2} + \frac{\partial^2 w}{\partial y \partial z} = \frac{1}{\mu} \frac{\partial p}{\partial z}. \qquad (8)$$

Now by making use of assumption 7 in the analysis, the second term on the left-hand side of equations 7 and 8 is omitted, and we have the reduced Navier-Stokes equations 9 and 10.

$$\frac{\partial^2 u}{\partial y^2} = \frac{1}{\mu} \frac{\partial p}{\partial x} \qquad (9)$$

$$\frac{\partial^2 w}{\partial y^2} = \frac{1}{\mu} \frac{\partial p}{\partial z}. \qquad (10)$$

In cylindrical coordinates, these equations become:

$$\frac{\partial^2 v_r}{\partial y^2} = \frac{1}{\mu} \frac{\partial p}{\partial r} \qquad (11)$$

$$\frac{\partial^2 v_\theta}{dy^2} = \frac{1}{\mu r} \frac{\partial p}{\partial \theta} \qquad (12)$$

where v_r and v_θ are the velocities in the r and θ directions, respectively.

REFERENCE

1. Sternlicht, B., and Pinkus, O. Theory of Hydrodynamic Lubrication, McGraw-Hill Book Company, Inc., 1961

B. AN INVESTIGATION OF PHOTOELECTRICALLY HETERODYNING COHERENT LIGHT by Cullen Q. Lee

In heterodyne detection of coherent light, the following conditions must be met: (1) The light beams must overlap on the photocathode surface, (2) the beams must be aligned almost parallel to each other, and (3) the beams must have the same direction of polarization. The heterodyne amplitude decreases exponentially with beam separation. The decrease in amplitude with beam misalignment depends on the field intensity distribution and detector width. For a Gaussian field intensity distribution and small detector width, heterodyning is possible over several minutes of beam misalignment. A Gaussian field intensity distribution gives better mixing for a given beam misalignment than a uniform field distribution. The difference in direction of polarization decreases heterodyne amplitude by the cosine of the angle between the directions of polarization. Since the cosine of angles near zero is quite broad, a slight difference in direction of polarization does not critically affect the heterodyne amplitude.

INTRODUCTION

The coherent light produced by a continuously operating laser is an ideal carrier for space communication. The high frequency of light permits transmission of a vast amount of intelligence on a single carrier. The extremely small divergence of the laser beam allows a large proportion of the transmitted energy to fall on the receiving device.

Transmission of intelligence on a laser beam is achieved much like transmission on a radio beam; the high frequency carrier is varied in amplitude or frequency at the rate of the signal being transmitted. Laser transmission consists of a light beam whose frequency or amplitude varies at the rate of the signal being transmitted. At the receiving station, the signal must be detected, or separated from the carrier. This may be done by mixing, or heterodyning, the received signal with a second signal of a slightly different frequency. The difference frequency is a modulated RF signal from which the transmitted intelligence may be retrieved.

The method of detection investigated here is that of heterodyning the received signal with a local signal, a beam from a second laser, by causing the two beams to overlap on a photomultiplier. The principal advantage of detection by heterodyning is an extremely high signal-to-noise ratio. Several questions arise as to the effects of varying the ways in which the two beams converge on the surface of the photomultiplier.

The field intensity distribution is investigated for two cases: (1) uniform intensity over the spot cast by the beam and (2) Gaussian distribution of intensity, that is, with an area of high intensity at the center declining exponentially to the edges.

Light beam divergence or lack of divergence determines whether the wave front presented to the photomultiplier will be spherical or plane. The effects of these two possibilities are investigated.

The light from a laser is polarized, much like ordinary light that has passed through a polarizing filter. The effect of alignment or misalignment of the polarization of the two laser beams on heterodyne current is determined.

Beam displacement, the amount of overlap of the two beams, is varied to determine the effect of such variations on the amount of heterodyne current produced.

Deviations in the parallelism of the two beams are introduced and the effects of these deviations on heterodyne current are determined.

The effects of rectangular and circular photomultiplier apertures are compared.

The investigation is conducted at two levels; theoretical and experimental. The theoretical investigation follows classical electromagnetic theory and assumes attainable conditions so that theory may be verified experimentally.

NOMENCLATURE

A	surface area
Å	Angstrom
a^2	$(\sigma_1^2 \cos^2 \theta_2 + \sigma_2^2 \cos^2 \theta_1)/4\,\sigma_1^2 \sigma_2^2$

b^2	$(\sigma_1^2 + \sigma_2^2)/4\,\sigma_1^2\,\sigma_2^2$
c	$m_x \cos\theta_2/2\sigma_2^2$
$\vec{E}$	electric field vector
()*	complex conjugate of ()
erf (x)	error function of x
exp x	exponential of x
g	constant
$\vec{H}$	magnetic field vector
h	$m_y/2\sigma_2^2$
I	integral
I_{ac}	alternating electric current
I_{dc}	direct electric current
i	$(-1)^{\frac{1}{2}}$
$(I_{ac})_R$	heterodyne current for rectangular aperture
$(I_{ac})_C$	heterodyne current for circular aperture
K	constant
k	$2\pi/\lambda$
ℓ	length
m_x	mean of Gaussian distribution in x-direction
m_y	mean of Gaussian distribution in y-direction
$\vec{n}$	unit normal vector
P	power
<P>	average power
p^2	$\dfrac{\sigma_1^2 \sin^2\theta_2 + \sigma_2^2 \sin^2\theta_1}{4\,\sigma_1^2\,\sigma_2^2}$
R	radius
R_o	radius of circular aperture
r	distance
Re $\vec{E}$	real part of $\vec{E}$
$\vec{S}$	$\vec{E} \times \vec{H}$ Poynting vector
t	time
w	detector aperture half-width
x, y, z	rectangular Cartesian coordinate axes
α	$2\pi\,(\sin\theta_2/\lambda_2 - \sin\theta_1/\lambda_1)$
β	phase angle
γ	$\dfrac{\pi\,\theta_o}{\sqrt{2}\,\sigma_2\,\lambda_2}$
Δx	increment of x
ϵ	permittivity
θ	angle
θ_0	laser beam divergence angle
θ_1	angle local oscillator makes with cathode surface normal
θ_2	angle signal beam makes with cathode surface normal
μ	permeability
ξ	$ax - (c \pm i\alpha)/2a$
σ	standard deviation
σ_1	standard deviation of local oscillator field intensity
σ_2	standard deviation of signal field intensity
φ	angle in polar coordinates
Ψ	arctan (γ/a^2)

ω	angular frequency
ω_1	angular frequency of local oscillator
ω_2	angular frequency of signal
Ω	angle between direction of polarization of local oscillator and signal

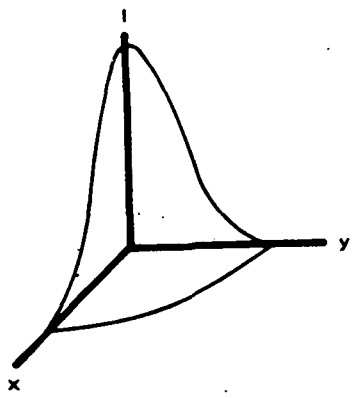


FIGURE 1. FIELD INTENSITY DISTRIBUTION OVER A NORMAL SURFACE

THEORETICAL INVESTIGATION

Two different types of field intensity distributions, a uniform distribution and Gaussian distribution, are used in deriving the equations.

A. DERIVATION OF PHOTOMIXING EQUATIONS FOR TWO LASER BEAMS USING RECTANGULAR DETECTOR APERTURE

The necessary conditions for photomixing two laser beams having plane wave fronts and for photomixing using a plane wave and a spherical wave are investigated.

1. Determination of Field Intensity Distribution. The current on a photocathode is given by the relation

$$I = K \int_A E^2 \, dA \tag{1}$$

where

K = constant

$\vec{E}$ = electric field vector

dA = element of surface area.

Before equation 1 can be evaluated, it is necessary to know the field intensity distribution. It has been shown theoretically (Refs. 1, 2, and 3) that the electric field intensity of a laser beam, when adjusted to operate in the fundamental mode of the transverse electromagnetic wave (TEM_{00}), will have a Gaussian distribution over a normal surface as shown in Figure 1.

The power contained in the laser beam may be determined by integrating the Poynting vector $\vec{S}$, for the field, over a surface normal to the beam. This is written as

$$P = \int_A \vec{S} \cdot \vec{n} \, dA \tag{2}$$

where

$$\vec{S} = \vec{E} \times \vec{H}. \tag{3}$$

The magnetic vector $\vec{H}$ for a plane electromagnetic wave is related to the electric vector $\vec{E}$ by

$$\vec{H} = \frac{\vec{k} \times \vec{E}}{\omega \mu}. \tag{4}$$

Thus,

$$\vec{S} = -\frac{1}{\omega \mu} \vec{E} \times (\vec{k} \times \vec{E}) = \frac{\vec{E} \cdot \vec{E}}{\omega \mu} \vec{k}. \tag{5}$$

For a surface normal to the direction of propagation

$$\vec{k} \cdot \vec{n} = k = \omega \sqrt{\mu \epsilon}. \tag{6}$$

The power given by equation 2 then becomes

$$P = \sqrt{\frac{\epsilon}{\mu}} \int_A \vec{E} \cdot \vec{E} \, dA. \tag{7}$$

The electric field over the surface may be written

$$\vec{E} = \mathrm{Re}\,\vec{E}_o \exp\{-i(\omega t - \vec{k}\cdot\vec{r}) - (x^2+y^2)/4\sigma^2\}$$
$$= \tfrac{1}{2}\exp\{-(x^2+y^2)/4\sigma^2\}\left[\vec{E}_o \exp\{-i(\omega t - \vec{k}\cdot\vec{r})\} + \vec{E}_o^{*} \exp\{i(\omega t - \vec{k}\cdot\vec{r})\}\right] \quad (8)$$

and

$$\vec{E}\cdot\vec{E} = \tfrac{1}{4}\exp\{-(x^2+y^2)/2\sigma^2\}\left[E_o^2 \exp\{-i2(\omega t - \vec{k}\cdot\vec{r})\} + 2\vec{E}_o\cdot\vec{E}_o^{*} + \{E_o^{*}\}^2 \exp\{i2(\omega t - \vec{k}\cdot\vec{r})\}\right] . \quad (9)$$

Substituting the time-independent term of equation 9 into equation 7 and integrating over an infinite plane perpendicular to the direction of propagation will give the average power contained in the laser beam. Equation 7 then becomes

$$\langle P\rangle = \tfrac{1}{2}\{\epsilon/\mu\}^{\frac{1}{2}}\vec{E}_o\cdot\vec{E}_o^{*}\int_{-\infty}^{\infty}\exp(-x^2/2\sigma^2)\,dx \int_{-\infty}^{\infty}\exp(-y^2/2\sigma^2)\,dy = \pi\sigma^2\{\epsilon/\mu\}^{\frac{1}{2}}E_o^2. \quad (10)$$

For free space,

$$\{\mu/\epsilon\}^{\frac{1}{2}} = 120\,\pi \text{ ohms}. \quad (11)$$

Then

$$E_o = (2/\sigma)\{30\langle P\rangle\}^{\frac{1}{2}}. \quad (12)$$

The standard deviation σ of the field intensity distribution may be determined experimentally by driving a photomultiplier, with a slit for an aperture, across the laser beam and measuring the dc current as a function of position. Substituting the time independent part of equation 9 into equation 1 gives

$$I_{dc} = \tfrac{1}{2}K\vec{E}_o\cdot\vec{E}_o^{*}\int_A \exp\{-(x^2+y^2)/2\sigma^2\}\,dA. \quad (13)$$

By making the slit width Δx very small and integrating from $-\infty$ to $+\infty$ in the y-direction,

$$I_{dc} = \tfrac{1}{2}K\vec{E}_o\cdot\vec{E}_o^{*}\Delta x\exp(-x^2/2\sigma^2)\int_{-\infty}^{\infty}\exp(-y^2/2\sigma^2)\,dy = \left(\frac{\pi}{2}\right)^{\frac{1}{2}}\sigma K\vec{E}_o\cdot\vec{E}_o^{*}\Delta x\exp(-x^2/2\sigma^2). \quad (14)$$

The shape of the curve given by equation 14 is determined by the constant σ^2 in the exponent. Therefore, it is possible to determine the standard deviation σ from a plot of the current given by equation 14.

2. Photomixing Equations for Plane Waves with a Gaussian Field Intensity Distribution. If we take E_1 and E_2 as the electric field vectors of the local oscillator and signal, respectively, the field intensity distribution in the x-direction over a normal surface will be as shown in Figure 2. The field intensity distribution in the y-direction will be as shown in Figure 3. Since the laser beams have annular symmetry, the curves in Figures 2 and 3 have the same shape, but they may have different separations in the x- and y-directions. The electric field vector $\vec{E}_1$ for the local oscillator may then be represented by the relation (Ref. 4)

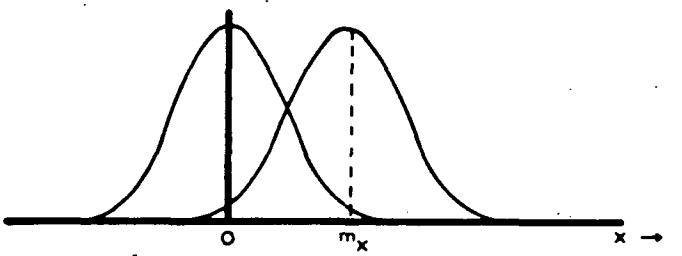


FIGURE 2. FIELD INTENSITY DISTRIBUTION FOR LASER IN X-DIRECTION

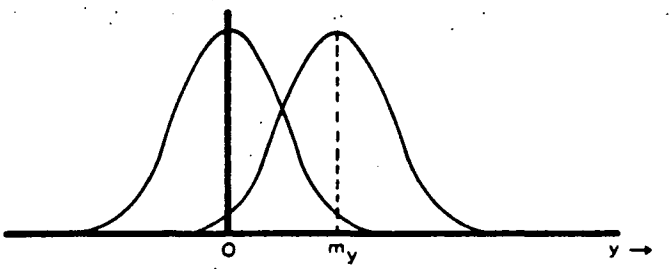


FIGURE 3. FIELD INTENSITY DISTRIBUTION FOR LASER IN Y-DIRECTION

$$\vec{E}_1 = \mathrm{Re}\,\vec{E}_{10}\exp\{-i(\omega_1 t - \vec{k}_1\cdot\vec{r}_1) - (x_1^2+y_1^2)/4\sigma_1^2\}$$
$$= \tfrac{1}{2}\exp\{-(x_1^2+y_1^2)/4\sigma_1^2\}\left[\vec{E}_{10}\exp\{-i(\omega_1 t - \vec{k}_1\cdot\vec{r}_1)\} + \vec{E}_{10}^{*}\exp\{i(\omega_1 t - \vec{k}_1\cdot\vec{r}_1)\}\right]. \quad (15)$$

The electric field vector $\vec{E}_2$ for the signal may be written

$$\vec{E}_2 = \tfrac{1}{2}\exp\{-[(x_2-m_x)^2+(y_2-m_y)^2]/4\sigma_2^2\}\left[\vec{E}_{20}\exp\{-i(\omega_2 t - \vec{k}_2\cdot\vec{r}_2)\} + \vec{E}_{20}^{*}\exp\{i(\omega_2 t - \vec{k}_2\cdot\vec{r}_2)\}\right]. \quad (16)$$

In equations 15 and 16, the coordinate system has been chosen so the xy-planes are normal to the direction of propagation. If the two laser beams are incident on a plane photocathode surface as shown in Figure 4, the coordinates used in equations 15 and 16 are related to those on the photocathode surface in the following manner:

$$x_1 = x \cos \theta_1 \tag{17}$$

$$y_1 = y_2 = y$$

$$x_2 = x \cos \theta_2.$$

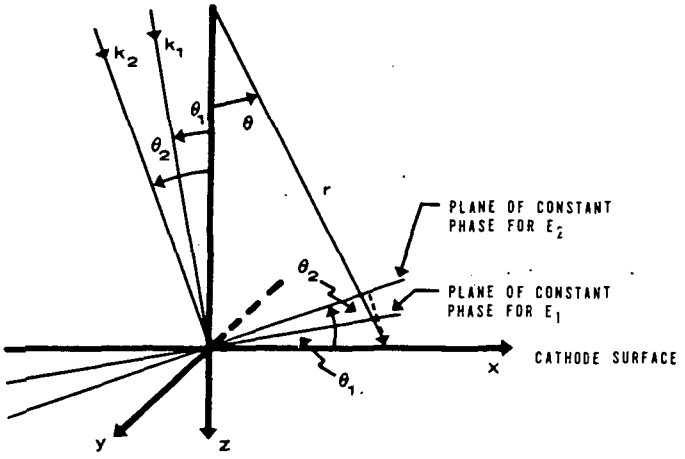


FIGURE 4. TWO LASER BEAMS INCIDENT ON PHOTOCATHODE SURFACE AT ANGLES θ_1 AND θ_2

In the new coordinate system, equations 15 and 16 become

$$\vec{E}_1 = \tfrac{1}{2}\exp\{-(x^2\cos^2\theta_1 + y^2)/4\sigma_1^2\} \left[\vec{E}_{10}\exp\{-i(\omega_1 t - \vec{k}_1\cdot\vec{r})\} + \vec{E}_{10}^{*}\exp\{i(\omega_1 t - \vec{k}_1\cdot\vec{r})\}\right] \tag{18}$$

and

$$\vec{E}_2 = \tfrac{1}{2}\exp\{-[(x\cos\theta_2 - m_x)^2 + (y-m_y)^2]/4\sigma_2^2\} \left[\vec{E}_{20}\exp\{-i(\omega_2 t - \vec{k}_2\cdot\vec{r})\} + \vec{E}_{20}^{*}\exp\{i(\omega_2 t - \vec{k}_2\cdot\vec{r})\}\right]. \tag{19}$$

The value of $\vec{E}$ to be inserted in equation 1 may now be determined.

$$E^2 = (\vec{E}_1 + \vec{E}_2)\cdot(\vec{E}_1 + \vec{E}_2) = E_1^2 + 2\vec{E}_1\cdot\vec{E}_2 + E_2^2. \tag{20}$$

From equations 18 and 19,

$$E_1^2 = \vec{E}_1\cdot\vec{E}_1 = \tfrac{1}{4}\exp\{-(x^2\cos^2\theta_1 + y^2)/2\sigma_1^2\} \left[E_{10}^2\exp\{-i2(\omega_1 t - \vec{k}_1\cdot\vec{r})\} + 2\vec{E}_{10}\cdot\vec{E}_{10}^{*} + \{E_{10}^{*}\}^2\exp\{i2(\omega_1 t - \vec{k}_1\cdot\vec{r})\}\right] \tag{21}$$

and

$$E_2^2 = \tfrac{1}{4}\exp\{-[(x\cos\theta_2 - m_x)^2 + (y-m_y)^2]/2\sigma_2^2\} \left[E_{20}^2\exp\{-i2(\omega_2 t - \vec{k}_2\cdot\vec{r})\} + 2\vec{E}_{20}\cdot\vec{E}_{20}^{*} + \{E_{20}^{*}\}^2\exp\{i2(\omega_2 t - \vec{k}_2\cdot\vec{r})\}\right]. \tag{22}$$

Also

$$2\vec{E}_1\cdot\vec{E}_2 = \tfrac{1}{2}\exp\{-(x^2\cos^2\theta_1 + y^2)/4\sigma_1^2 - [(x\cos\theta_2 - m_x)^2 + (y-m_y)^2]/4\sigma_2^2\} \left[\vec{E}_{10}\cdot\vec{E}_{20}\exp\{-i(\omega_2+\omega_1)t + i(\vec{k}_2+\vec{k}_1)\cdot\vec{r}\} + \vec{E}_{10}^{*}\cdot\vec{E}_{20}^{*}\exp\{i(\omega_2+\omega_1)t - i(\vec{k}_2+\vec{k}_1)\cdot\vec{r}\} + \vec{E}_{10}\cdot\vec{E}_{20}^{*}\exp\{i(\omega_2-\omega_1)t - i(\vec{k}_2-\vec{k}_1)\cdot\vec{r}\} + \vec{E}_{10}^{*}\cdot\vec{E}_{20}\exp\{-i(\omega_2-\omega_1)t + i(\vec{k}_2-\vec{k}_1)\cdot\vec{r}\}\right]. \tag{23}$$

The terms in equations 21, 22, and 23 containing optical-frequency exponents will produce no heterodyne current in the detector and therefore may be neglected. The dc terms in equations 21 and 23 may also be neglected. The beat-frequency current is then given by

$$I_{ac} = \tfrac{1}{2}K\vec{E}_{10}\cdot\vec{E}_{20}^{*}\exp\{-(m_x^2+m_y^2)/4\sigma_2^2\}\exp\{i(\omega t-\beta)\} \int_{-\ell}^{\ell}\exp(-b^2y^2+hy)\,dy\int_{-w}^{}\exp\{-a^2x^2+(c-i\alpha)x\}\,dx + \tfrac{1}{2}K\vec{E}_{10}^{*}\cdot\vec{E}_{20}\exp\{-(m_x^2+m_y^2)/4\sigma_2^2\}\exp\{-i(\omega t-\beta)\} \int_{-\ell}^{\ell}\exp(-b^2y^2+hy)\,dy\int_{-w}^{w}\exp\{-a^2x^2+(c+i\alpha)x\}dx \tag{24}$$

where the new parameters are defined by

$$\omega = \omega_2 - \omega_1$$

$$(\vec{k_2} - \vec{k_1}) \cdot \vec{r} = \beta + \alpha x$$

$$x = r \sin\theta$$

$$\alpha = 2\pi \left(\sin\theta_2/\lambda_2 - \sin\theta_1/\lambda_1\right)$$

$$\beta = 2\pi r \cos\theta \left(\cos\theta_2/\lambda_2 - \cos\theta_1/\lambda_1\right) \qquad (25)$$

$$a^2 = (\sigma_1^2 \cos^2\theta_2 + \sigma_2^2 \cos^2\theta_1)/4\,\sigma_1^2\sigma_2^2$$

$$b^2 = (\sigma_1^2 + \sigma_2^2)/4\sigma_1^2\sigma_2^2$$

$$c = m_x \cos\theta_2/2\sigma_2^2$$

$$h = m_y/2\sigma_2^2 .$$

In equation 24 one type of integral must be evaluated, i. e.,

$$I = \int_{-w}^{w} \exp\{-a^2x^2 + (c \pm i\alpha)x\}\,dx. \qquad (26)$$

This integral cannot be evaluated in closed form between finite limits, making numerical integration necessary. To make equation 26 easier to program for a computer, it may be written as

$$I = \int_{-w}^{w} \exp(-a^2x^2 + cx)\cos\alpha x\,dx \pm i\int_{-w}^{w} \exp(-a^2x^2 + cx)\sin\alpha x\,dx. \qquad (27)$$

For computer programing,

$$I_r = \int_{-w}^{w} \exp(-a^2x^2 + cx)\cos\alpha x\,dx \qquad (28)$$

$$I_i = \int_{-w}^{w} \exp(-a^2x^2 + cx)\sin\alpha x\,dx \qquad (29)$$

and

$$I = \{I_r^2 + I_i^2\}^{\frac{1}{2}}. \qquad (30)$$

To gain insight into the effect of beam separation on photoheterodyning, take the case where the cathode surface is much larger than the laser beam diameter. Under these conditions the limits of integration may be taken from $-\infty$ to $+\infty$. Equation 26 may now be written as

$$I = \exp\{(c \pm i\alpha)^2/4a^2\}\int_{-\infty}^{\infty} \exp\{-[ax - (c \pm i\alpha)/2a]^2\}\,dx. \qquad (31)$$

If the change of variables is made,

$$\xi = ax - (c \pm i\alpha)/2a, \qquad (32)$$

equation 31 may be written as

$$I = (1/a)\exp[(c \pm i\alpha)^2/4a^2]\int_{-\infty \pm i\alpha/2a}^{\infty \pm i\alpha/2a} \exp(-\xi^2)\,d\xi. \qquad (33)$$

The contour of integration is $\alpha/2a$ units above or below the real axis from $-\infty$ to $+\infty$. Since there are no poles between this line and the real axis, we may take the real integral

$$I = (1/a)\exp[(c \pm i\alpha)^2/4a^2]\int_{-\infty}^{\infty} \exp(-\xi^2)\,d\xi$$

$$= (\sqrt{\pi}/a)\exp[(c \pm i\alpha)^2/4a^2]. \qquad (34)$$

Also from equation 24,

$$\int_{-\infty}^{\infty} \exp(-b^2y^2 + hy)\,dy = (\sqrt{\pi}/b)\exp(h^2/4b^2). \qquad (35)$$

Substituting equations 34 and 35 into equation 24, the total current at the heterodyne frequency ω is obtained. Thus,

$$I_{ac} = (\pi/2ab)\,K\exp(-\alpha^2/4a^2)\exp\left[-\frac{m_x^2\cos^2\theta_1}{4(\sigma_1^2\cos^2\theta_2 + \sigma_2^2\cos^2\theta_1)} - \frac{m_y^2}{4(\sigma_1^2 + \sigma_2^2)}\right]\left[\vec{E}_{10}\cdot\vec{E}_{20}^{\,*}\exp\{i(\omega t - \beta - c\alpha/2a^2)\} + \vec{E}_{10}^{\,*}\cdot\vec{E}_{20}\exp\{-i(\omega t - \beta - c\alpha/2a^2)\}\right]. \qquad (36)$$

From equation 36 it is seen that the heterodyne current decreases exponentially with beam separation and beam misalignment angle. It also depends on the polarization. For linear polarization, $\vec{E}_{10}$ and $\vec{E}_{20}$ are real so equation 36 may be written as

$$I_{ac} = (\pi/ab)\, K\vec{E}_{10}\cdot\vec{E}_{20}$$

$$\exp\left[-\frac{m_x^2\cos^2\theta_1}{4(\sigma_1^2\cos^2\theta_2+\sigma_2^2\cos^2\theta_1)}-\frac{m_y^2}{4(\sigma_1^2+\sigma_2^2)}\right]$$

$$\exp(-\alpha^2/4a^2)\cos(\omega t-\beta-c\alpha/2a^2). \qquad (37)$$

From equation 37 it is seen that when the $\vec{E}$ vectors are oriented 90 degrees to each other no heterodyning will occur.

To determine the effect of detector aperture on heterodyning, examine the case when the two beams are coincident on the photocathode. Under these conditions,

$$m_x = m_y = c = h = 0$$

and equation 24 becomes

$$I_{ac} = K\vec{E}_{10}\cdot\vec{E}_{20}\int_{-\ell}^{\ell}\exp(-b^2y^2)\,dy$$

$$\int_{-w}^{w}\exp(-a^2x^2)\cos(\omega t-\beta-\alpha x)\,dx. \qquad (38)$$

The cosine term in equation 38 may be written

$$\cos(\omega t-\beta-\alpha x) = \cos(\omega t-\beta)\cos\alpha x + \sin(\omega t-\beta)\sin\alpha x$$

but

$$\int_{-w}^{w}\exp(-a^2x^2)\sin\alpha x\,dx = 0$$

since the integrand is an odd function. Then equation 38 becomes

$$I_{ac} = K\vec{E}_{10}\cdot\vec{E}_{20}\cos(\omega t-\beta)\int_{-\ell}^{\ell}\exp(-b^2y^2)\,dy$$

$$\int_{-w}^{w}\exp(-a^2x^2)\cos\alpha x\,dx. \qquad (39)$$

The first integral on the right of equation 39 has its maximum value when $\ell\to\infty$. However, this is not true for the second integral. To determine what values of w will give maximum heterodyne current, take

$$I = \int_{-w}^{w}\exp(-a^2x^2)\cos\alpha x\,dx \qquad (40)$$

and differentiate under the integral sign (Ref. 5) with respect to w obtaining

$$\frac{\partial I}{\partial w} = 2\exp(-a^2w^2)\cos\alpha w. \qquad (41)$$

Then

$$\partial I/\partial w = 0$$

when

$$\alpha w = (2n-1)\frac{\pi}{2}. \qquad (42)$$

The first maximum will give the largest value of I so the optimum detector aperture is given by

$$w = \frac{\pi}{2\alpha}\ \text{mm}, \qquad (43)$$

remembering that w is the detector aperture half-width. As an example, suppose that an angle of 20 arc seconds is between the local oscillator and the incoming signal. At a wavelength of 6328Å, this would make $\alpha \approx 1$. Then,

$$w = \frac{\pi}{2}\ \text{mm},$$

or the total width of the detector aperture should be

$$2w = \pi\ \text{mm}. \qquad (44)$$

Further insight into the effect of parameter variation on heterodyning may be gained by holding w fixed and differentiating equation 40 with respect to α. Thus,

$$\frac{\partial I}{\partial\alpha} = -\int_{-w}^{w} x\exp(-a^2x^2)\sin\alpha x\,dx. \qquad (45)$$

After an integration by parts, equation 45 becomes

$$\frac{\partial I}{\partial\alpha} = (1/a^2)\exp(-a^2w^2)\sin\alpha w - \alpha I/2a^2. \qquad (46)$$

Setting $\partial I/\partial\alpha = 0$ gives the maximum value of I as

$$I_m = (2/\alpha)\exp(-a^2w^2)\sin\alpha w. \qquad (47)$$

Holding w and α fixed and differentiating equation 40 with respect to the parameter a will give

$$\frac{\partial I}{\partial a} = -2a \int_{-w}^{w} x^2 \exp(-a^2 x^2) \cos \alpha x \, dx. \quad (48)$$

Integrating twice by parts gives

$$\frac{\partial I}{\partial a} = (2w/a) \exp(-a^2 w^2) \cos \alpha w + (\alpha^2/2a^3 - 1/a) I - (\alpha/a^3) \exp(-a^2 w^2) \sin \alpha w. \quad (49)$$

If equation 49 is set equal to zero and solved for I, the integral of equation 40 will be maximized with respect to a. Thus,

$$I = \frac{2\alpha}{\alpha^2 - 2a^2} \exp(-a^2 w^2) \sin \alpha w - \frac{4a^2 w}{\alpha^2 - 2a^2} \exp(-a^2 w^2) \cos \alpha w. \quad (50)$$

Differentiating equation 50 with respect to α and setting $\partial I / \partial \alpha = 0$,

$$\tan \alpha w = \frac{\alpha^3 w + 2a^2 \alpha w}{\alpha^2 + 2a^2 + 4a^4 w^2 - 2a^2 w^2 \alpha^2}. \quad (51)$$

The roots of equation 51 determine the values of α that will make I a maximum or a minimum for given values of w and a.

3. Photomixing Equations for Plane Waves with a Uniform Field-Intensity Distribution. From equation 24, the ac term for the current may be written as

$$I_{ac} = \tfrac{1}{2} K \vec{E}_{10} \cdot \vec{E}_{20}^{*} \exp\{i(\omega t - \beta)\} \int_{-\ell}^{\ell} \int_{-w}^{w} \exp(-i\alpha x) \, dy \, dx + \tfrac{1}{2} K \vec{E}_{10}^{*} \cdot \vec{E}_{20} \exp\{-i(\omega t - \beta)\} \int_{-\ell}^{\ell} \int_{-w}^{w} \exp(i\alpha x) \, dy \, dx. \quad (52)$$

This integrates to give

$$I_{ac} = 2w\ell K \left[\vec{E}_{10} \cdot \vec{E}_{20}^{*} \exp\{i(\omega t - \beta)\} + \vec{E}_{10}^{*} \cdot \vec{E}_{20} \exp\{-i(\omega t - \beta)\} \right] \frac{\sin \alpha w}{\alpha w}. \quad (53)$$

For linear polarization, equation 53 becomes

$$I_{ac} = 4w\ell K \vec{E}_{10} \cdot \vec{E}_{20} \cos(\omega t - \beta) \left[\frac{\sin \alpha w}{\alpha w} \right]. \quad (54)$$

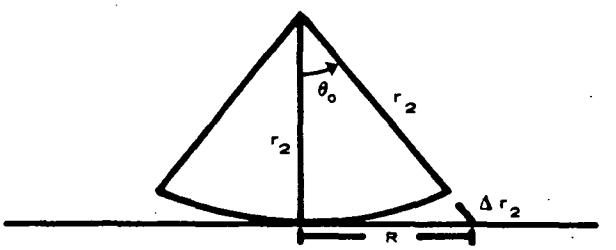


FIGURE 5. PLANE LASER BEAM MIXED WITH DIVERGING LASER BEAM

4. Photomixing a Plane Laser Beam with a Diverging Laser Beam. To determine the effect of beam divergence on photoelectrical heterodyning, take the case of one parallel beam being mixed with a diverging beam. Also let the two beams be coincident on a plane photocathode surface with the central ray of the diverging beam parallel to the other beam and normal to the photocathode surface as shown in Figure 5. Thus,

$$R^2 + r_2^2 = (r_2 + \Delta r_2)^2 = r_2^2 + 2r_2 \Delta r_2 + (\Delta r_2)^2$$

and for $(\Delta r_2)^2 \ll 2r_2$,

$$\Delta r_2 \approx R^2 / 2r_2. \quad (55)$$

The phase factor to be used in equation 24 may now be written as

$$(\vec{k}_2 - \vec{k}_1) \cdot \vec{r} = \beta + k_2 \Delta r_2 \approx \beta + \frac{\pi R^2}{r_2 \lambda_2} \quad (56)$$

where R is the radial distance from the central ray on the photocathode surface. The value of r_2 is determined from the beam divergence angle θ_0 and the standard deviation σ_2 of the Gaussian field intensity distribution as shown in Figure 6. The beam divergence angle is taken between the central ray and the ray where the field intensity has decreased by the factor e^{-1}. From Figure 6,

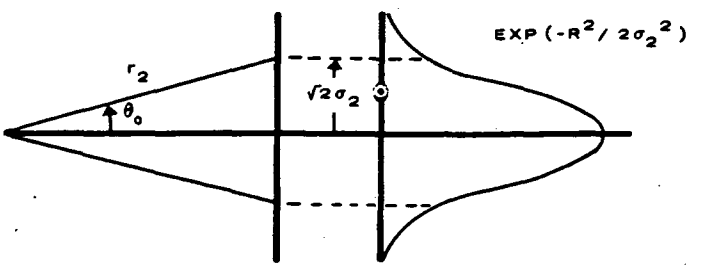


FIGURE 6. RELATION BETWEEN BEAM DIVERGENCE ANGLE AND STANDARD DEVIATION

$$\theta_0 \approx \sin\theta_0 = \frac{\sqrt{2}\,\sigma_2}{r_2}$$

or

$$r_2 \approx \sqrt{2}\,\sigma_2/\theta_0 . \tag{57}$$

Equation 56 may then be written as

$$(\vec{k}_2 - \vec{k}_1)\cdot\vec{r} = \beta + \gamma R^2 \tag{58}$$

where

$$\gamma = \frac{\pi\,\theta_0}{\sqrt{2}\,\sigma_2\,\lambda_2} . \tag{59}$$

Substituting equation 58 into equation 24 and taking the case of linear polarization,

$$I_{ac} = K\vec{E}_{10}\cdot\vec{E}_{20}\int_A \exp(-a^2R^2)\cos(\omega t-\beta-\gamma R^2)\,dA. \tag{60}$$

For a circular aperture,

$$dA = 2\pi\,RdR. \tag{61}$$

If a change of variables is made and the value of area element given by equation 61 is used, equation 60 may be written as

$$I_{ac} = \pi K\vec{E}_{10}\cdot\vec{E}_{20}\cos(\omega t-\beta)\int_0^{R_o^2}\exp(-a^2\xi)\cos\gamma\xi\,d\xi$$

$$+\pi K\vec{E}_{10}\cdot\vec{E}_{20}\sin(\omega t-\beta)\int_0^{R_o^2}\exp(-a^2\xi)\sin\gamma\xi\,d\xi \tag{62}$$

where R_o is the radius of the detector aperture. Equation 62 integrates to give

$$I_{ac} = \frac{\pi K\,\vec{E}_{10}\cdot\vec{E}_{20}}{a^4+\gamma^2}\Big[\cos(\omega t-\beta)\{a^2 + \exp(a^2R_o^2)$$

$$[\gamma\sin(\gamma R_o^2) - a^2\cos(\gamma R_o^2)]\} + \sin(\omega t-\beta)\{\gamma$$

$$-\exp(-a^2R_o^2)\,[a^2\sin(\gamma R_o^2) + \gamma\cos(\gamma R_o^2)]\}\Big] . \tag{63}$$

For the case where the photocathode surface is larger than the light beam, let $R_o \to \infty$ so that equation 63 becomes

$$I_{ac} = \frac{\pi K\,\vec{E}_{10}\cdot\vec{E}_{20}}{\{a^4+\gamma^2\}^{\frac{1}{2}}}\cos(\omega t-\beta-\psi) \tag{64}$$

where

$$\psi = \arctan(\gamma/a^2). \tag{65}$$

Equation 64 reduces to the plane wave case when $\gamma = 0$. It is seen that the beam divergence reduces the heterodyne current amplitude by the factor,

$$\frac{a^2}{\{a^4+\gamma^2\}^{\frac{1}{2}}}.$$

B. DERIVATION OF PHOTOMIXING EQUATIONS FOR TWO LASER BEAMS USING CIRCULAR APERTURE

To compare photomixing using a circular aperture with that for a rectangular aperture, take the case when the local oscillator beam and signal beam are coincident on the photocathode surface. The electric field vector $\vec{E}_1$ for the local oscillator may then be written,

$$\vec{E}_1 = \tfrac{1}{2}\exp\{-R^2(1-\sin^2\theta_1\cos^2\varphi)/4\sigma_1^2\}$$

$$\Big[\vec{E}_{10}\exp\{-i(\omega_1 t-\vec{k}_1\cdot\vec{r})\} + \vec{E}_{10}^{*}\exp\{i(\omega_1 t-\vec{k}_1\cdot\vec{r})\}\Big]. \tag{66}$$

The electric field vector $\vec{E}_2$ for the signal is

$$\vec{E}_2 = \tfrac{1}{2}\exp\{-R^2(1-\sin^2\theta_2\cos^2\varphi)/4\sigma_2^2\}$$

$$\Big[\vec{E}_{20}\exp\{-i(\omega_2 t-\vec{k}_2\cdot\vec{r})\} + \vec{E}_{20}^{*}\exp\{i(\omega_2 t-\vec{k}_2\cdot\vec{r})\}\Big]. \tag{67}$$

The product terms of equations 66 and 67 that produce heterodyne current are

$$2\vec{E}_1\cdot\vec{E}_2 = \tfrac{1}{2}\exp\{-(\sigma_1^2+\sigma_2^2)\,R^2/4\,\sigma_1^2\,\sigma_2^2\}$$

$$\exp\{(\sigma_1^2\sin^2\theta_2 + \sigma_2^2\sin^2\theta_1)\cos^2\varphi\,R^2/4\sigma_1^2\,\sigma_2^2\}$$

$$\Big[\vec{E}_{10}\cdot\vec{E}_{20}^{*}\exp\{i(\omega_2-\omega_1)t - i(\vec{k}_2-\vec{k}_1)\cdot\vec{r}\}$$

$$+\vec{E}_{10}^{*}\cdot\vec{E}_{20}\exp\{-i(\omega_2-\omega_1)t + i(\vec{k}_2-\vec{k}_1)\cdot\vec{r}\}\Big]. \tag{68}$$

Substituting equation 68 in equation 1 gives

$$I_{ac} = \tfrac{1}{2} K\vec{E}_{10} \cdot \vec{E}_{20}^{*} \int_0^{R_o} \int_0^{2\pi} \exp\{(p^2 \cos^2\varphi - b^2)R^2 + i(\omega t - \beta - \alpha R\cos\varphi)\}\, RdRd\varphi + \tfrac{1}{2} K\vec{E}_{10}^{*} \cdot \vec{E}_{20} \int_0^{R_o} \int_0^{2\pi} \exp\{(p^2\cos^2\varphi - b^2)R^2 - i(\omega t - \beta - \alpha R\cos\varphi)\}\, RdRd\varphi \quad (69)$$

where

$$dA = RdRd\varphi$$

$$(\vec{k}_2 - \vec{k}_1) \cdot \vec{r} = \beta + \alpha R\cos\varphi$$

R_o = radius of aperture (70)

$$p^2 = \frac{\sigma_1^2 \sin^2\theta_2 + \sigma_2^2 \sin^2\theta_1}{4\sigma_1^2\sigma_2^2}$$

and all other symbols are as defined in equations 25. Equation 69 may be integrated with respect to φ obtaining

$$I_{ac} = \sqrt{\pi}\, K\left[\vec{E}_{10} \cdot \vec{E}_{20}^{*} \exp\{i(\omega t - \beta)\} + \vec{E}_{10}^{*} \cdot \vec{E}_{20} \exp\{-i(\omega t - \beta)\}\right] \sum_{n=0}^{\infty} \frac{(-1)^n \Gamma(n+\frac{1}{2})\, 2^n p^{2n}}{\alpha^n n!} \int_0^{R_o} \exp(-a^2R^2)\, J_n(\alpha R)\, R^{n+1}\, dR. \quad (71)$$

For linear polarization $\vec{E}_{10}$ and $\vec{E}_{20}$ are real; for normal incidence $p^2 \to 0$ so that, under these conditions, equation 71 may be written as

$$I_{ac} = 2\pi K \vec{E}_{10} \cdot \vec{E}_{20} \cos(\omega t - \beta) \int_0^{R_o} \exp(-b^2R^2)\, J_0(\alpha R)\, RdR. \quad (72)$$

The integrals in equations 71 and 72 cannot be evaluated in closed form for finite limits, but they are easily programed for computer evaluation.

To determine what value of R_o will give maximum heterodyne current, take

$$I = \int_0^{R_o} \exp(-a^2R^2) J_0(\alpha R)\, RdR \quad (73)$$

and differentiate under the integral sign with respect to R_o obtaining

$$\frac{\partial I}{\partial R_o} = \exp(-a^2R_o^2)\, J_0(\alpha R_o)\, R_o = 0. \quad (74)$$

The maximum current will then occur when

$$J_0(\alpha R_o) = 0 \quad (75)$$

or

$$R_o = 2.4/\alpha. \quad (76)$$

C. DISCUSSION OF THEORETICAL RESULTS

By examining some of the equations derived in the preceding sections, it is possible to ascertain optimum physical conditions for heterodyning two coherent light beams.

To determine the relative amplitudes of heterodyne current for the two aperture shapes examined, the equations for the current may be written in terms of the power contained in each laser beam. For a rectangular aperture,

$$(I_{ac})_R = \frac{K\sqrt{\mu/\epsilon}\sqrt{\langle P_1\rangle\langle P_2\rangle}}{\sqrt{\pi} b\, \sigma_1\, \sigma_2} \cos\Omega \cos(\omega t - \beta)\, \mathrm{erf}(b\ell) \int_{-w}^{w} \exp(-a^2x^2) \cos(\alpha x)\, dx, \quad (77)$$

and for a circular aperture

$$(I_{ac})_C = \frac{2K\sqrt{\mu/\epsilon}\sqrt{\langle P_1\rangle\langle P_2\rangle}}{\sigma_1\, \sigma_2} \cos\Omega \cos(\omega t - \beta) \int_0^{R_o} \exp(-a^2R^2)\, J_0(\alpha R)\, RdR. \quad (78)$$

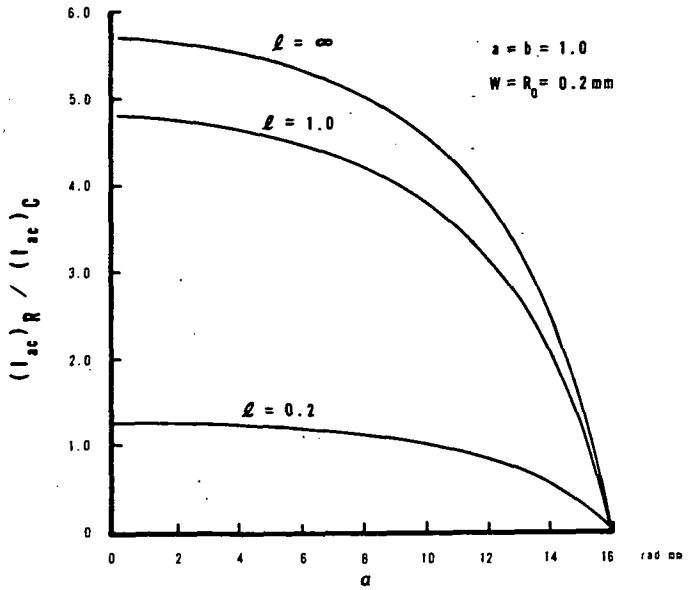


FIGURE 7. RATIO OF HETERODYNE CURRENT FOR RECTANGULAR APERTURE TO HETERODYNE CURRENT FOR CIRCULAR APERTURE

Taking the ratio of equation 77 to equation 78,

$$\frac{(I_{ac})_R}{(I_{ac})_C} = \frac{\operatorname{erf}(b\ell)}{2b\sqrt{\pi}} \frac{\int_{-w}^{w} \exp(-a^2x^2)\cos(\alpha x)\,dx}{\int_0^{R_o} \exp(-a^2R^2)\,J_o(\alpha R)\,R\,dR} \qquad (79)$$

This ratio is plotted in Figure 7 versus α for a=b=1.0 and $w=R_o=0.2$ mm using three different slit lengths 2ℓ. These curves show that for small values of α a rectangular aperture with $\ell \gg w$ will give larger values of heterodyne current than a circular aperture having a diameter equal to the total slit width. However, the current decreases more rapidly with an increase in misalignment for the rectangular aperture than for a circular aperture.

In Figure 8 the normalized amplitude of heterodyne current is plotted versus α for a rectangular aperture using detector half-widths of 0.4 mm, 1 mm, and infinity. Figure 9 shows the normalized amplitude of heterodyne current versus α for a circular aperture using detector radii of 0.1 mm, 0.4 mm, and 1 mm. Several different field intensity distributions are shown

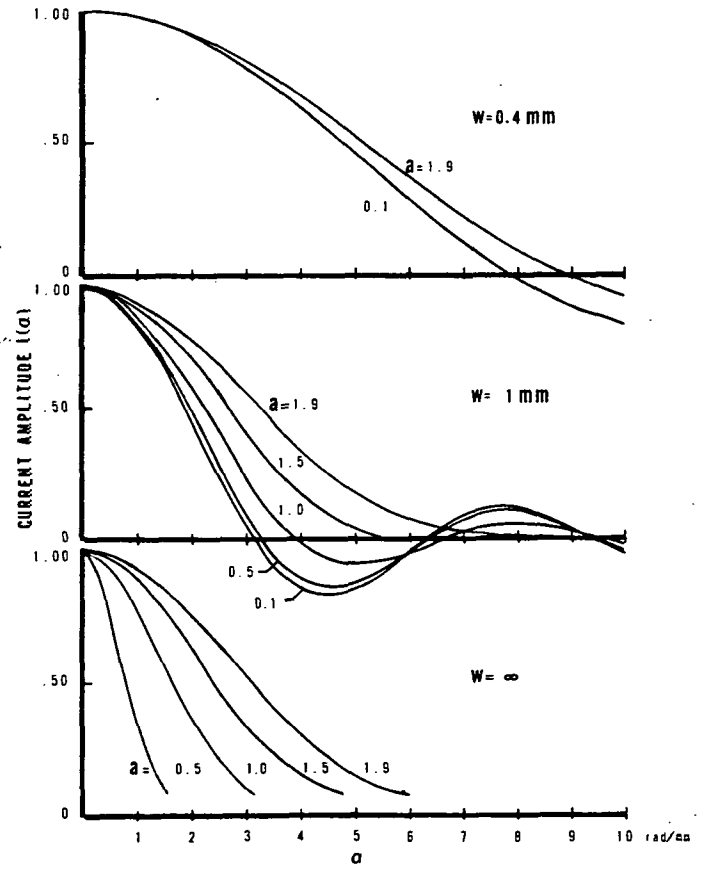


FIGURE 8. NORMALIZED HETERODYNE CURRENT VERSUS α FOR RECTANGULAR APERTURE

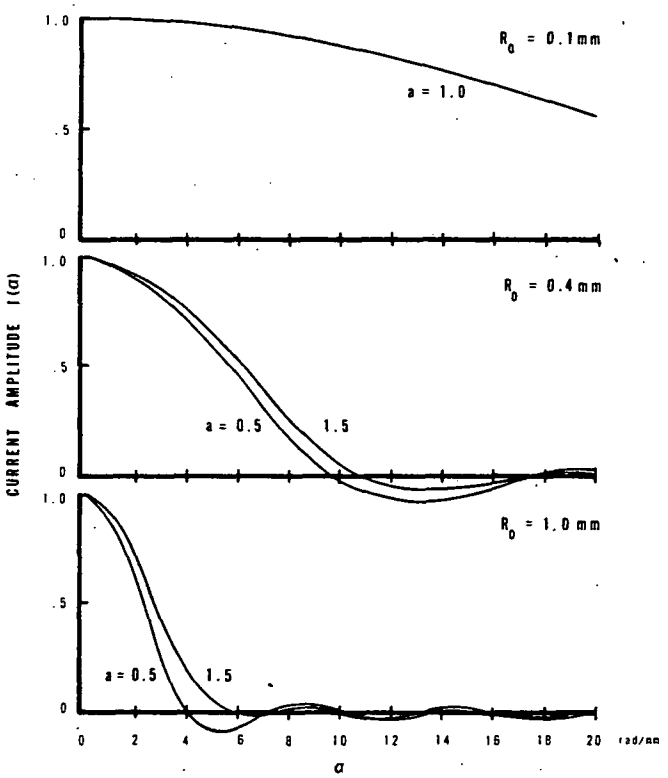


FIGURE 9. NORMALIZED HETERODYNE CURRENT VERSUS α FOR CIRCULAR APERTURE

for each detector size. For a given detector size, heterodyning will occur over larger values of misalignment for large values of a, since α is related to the angle between the two beams by the fourth of equations 25 and a is related to the Gaussian field distribution by the sixth of equations 25. In other words, if it is desired to heterodyne two light beams having an angle between them, the beams should be collimated down to produce a very small spot on the detector photocathode.

Figure 10 shows the normalized heterodyne current as a function of detector half-width for several values of α using a rectangular aperture. Maximum current occurs when $2w=\pi/\alpha$ as given by equation 43 for any given value of a. If α is practically zero, the detector width should be large enough to intercept most of the beam; but if α is appreciable, the detector size should be restricted by an aperture of width $2w=\pi/\alpha$.

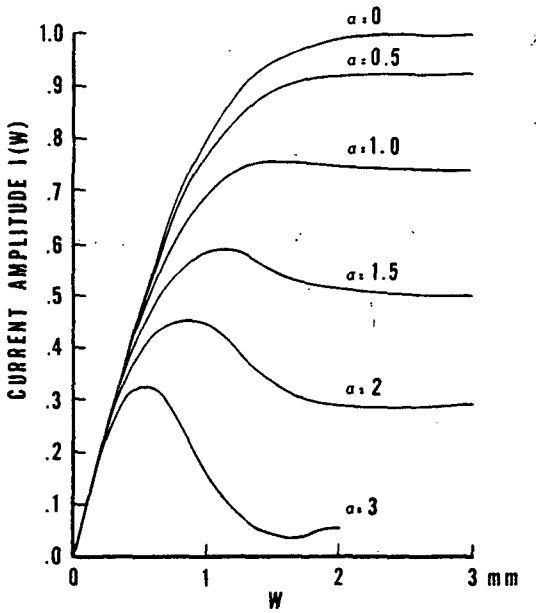


FIGURE 10. NORMALIZED HETERODYNE CURRENT VERSUS W FOR RECTANGULAR APERTURE

Equation 54 gives the heterodyne current for plane waves having a uniform field intensity distribution. If this curve was plotted in Figure 8, it would almost coincide with the curve shown for a=0.1.

Equation 64 gives the heterodyne current for a parallel beam having a plane wave front being mixed with a diverging beam having a spherical wave front. It is assumed that the two beams intercept the plane cathode surface at the same spot and that they are parallel to each other. Comparing this equation with equation 37 for two plane waves when $m_x=m_y=\alpha=0$, it is seen that the heterodyne current is reduced by the factor

$$\frac{a^2}{\{a^4+\gamma^2\}^{\frac{1}{2}}}$$

when a diverging beam is mixed with a parallel beam. The value of γ as given by equation 59 depends on the difference of curvature of the two beams.

It is also desirable to determine the effect on heterodyne current of collimating or focusing one beam down to a very small spot and leaving the other beam comparatively large. This may be done with the aid of equation 78. First, assume that the two beams intercept the cathode surface almost normally. Then

$$a^2=\frac{\sigma_1^2+\sigma_2^2}{4\sigma_1^2\sigma_2^2}. \tag{80}$$

If $\sigma_1 = g\sigma_2$, where g is some constant greater than unity, equation 80 may be written as

$$a^2=\frac{g^2+1}{4\sigma_1^2}. \tag{81}$$

The ratio of heterodyne current $(I_{ac})_=$ for $\sigma_1=\sigma_2$ to the heterodyne current $(I_{ac})_{\neq}$ for $\sigma_1 = g\sigma_2$ is

$$\frac{(I_{ac})_=}{(I_{ac})_{\neq}}=\frac{1}{g}\frac{\int_0^{R_o}\exp(-R^2/2\sigma^2)\,J_o(\alpha R)\,RdR}{\int_0^{R_o}\exp\left[-\frac{(g^2+1)R^2}{4\sigma^2}\right]J_o(\alpha R)\,RdR} \tag{82}$$

where σ has been written for σ_1. For the case when $\alpha = 0$, these integrals may be evaluated to give

$$\frac{(I_{ac})_=}{(I_{ac})_{\neq}}=\frac{g^2+1}{2g}\,\frac{\left[1-\exp(-R_o^2/2\sigma^2)\right]}{\left[1-\exp\{-(g^2+1)R_o^2\}\right]}. \tag{83}$$

Now let $R_o \to \infty$ and we have

$$\frac{(I_{ac})_=}{(I_{ac})_{\neq}}=\frac{g^2+1}{2g}. \tag{84}$$

This ratio is never greater than unity. Hence, for perfect alignment of the two beams and large detector

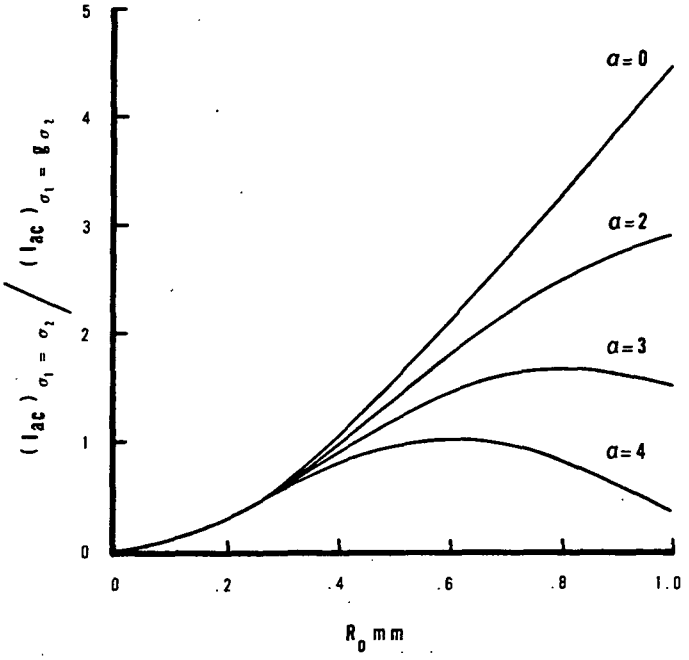


FIGURE 11. RATIO OF HETERODYNE CURRENT FOR EQUAL AND UNEQUAL σ's VERSUS R_0

apertures, best mixing will occur when the two beams have the same field intensity distribution. To show how unequal field intensity distributions affect heterodyning, the ratio given in equation 82 is plotted in Figure 11 versus aperture radius for several values of α using $\sigma_1 = 1/\sqrt{2}$ mm and $\sigma_2 = 0.05$ mm. From these curves it is seen that for $R_0 < 0.4$ mm collimating or focusing one beam down to a standard deviation of 0.05 mm will give more heterodyne current for all values of α. For $\alpha > 4$, improvement is indicated for all values of R_0. Figure 12 is a three-dimensional representation of normalized heterodyne current as a function of α and w for a = 1.0. Thus one can visualize the way in which the current amplitude depends on both the detector aperture size and the lack of alignment of the two beams being mixed.

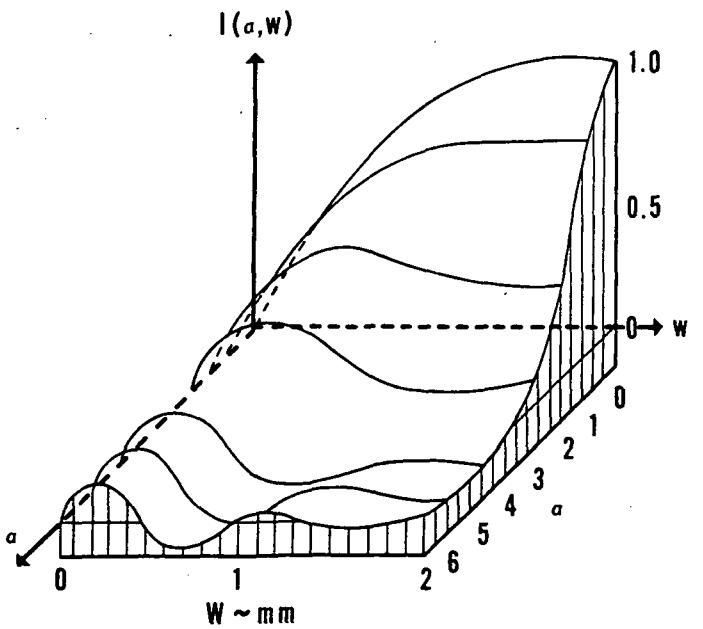


FIGURE 12. THREE-DIMENSIONAL REPRESENTATION OF NORMALIZED HETERODYNE CURRENT

COMPARISON OF EXPERIMENTAL AND THEORETICAL RESULTS

Experimental values of normalized heterodyne current as a function of detector half-width for rectangular apertures are shown in Figure 13. Comparing

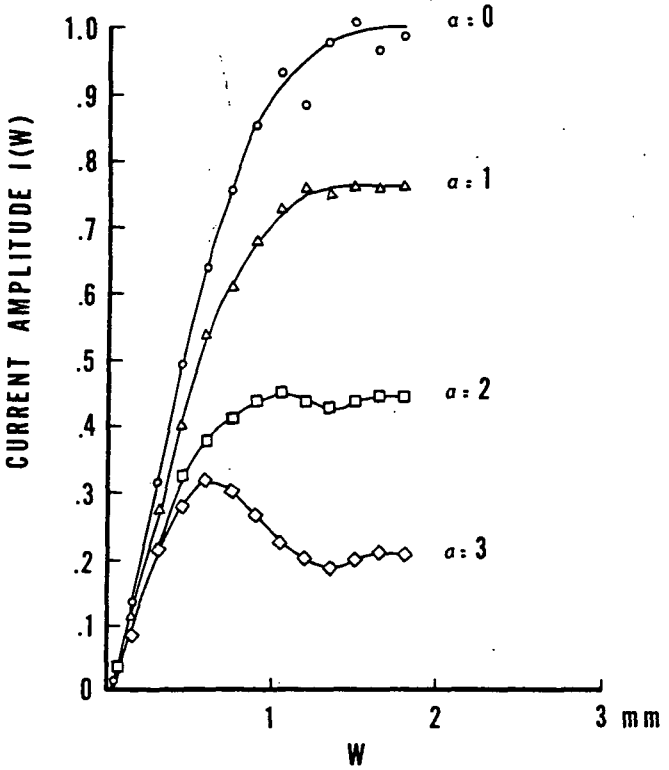


FIGURE 13. EXPERIMENTAL VALUES OF NORMALIZED HETERODYNE CURRENT

these curves with those shown in Figure 10 for the theoretical values shows very good agreement. Theoretical and experimental values of heterodyne current amplitude $I(\alpha)$ versus angle α are shown in Figures 14, 15, and 16 for rectangular apertures using several values of detector widths. The solid curves give theoretical values and the circled points give experimental values. Figure 17 shows theoretical and

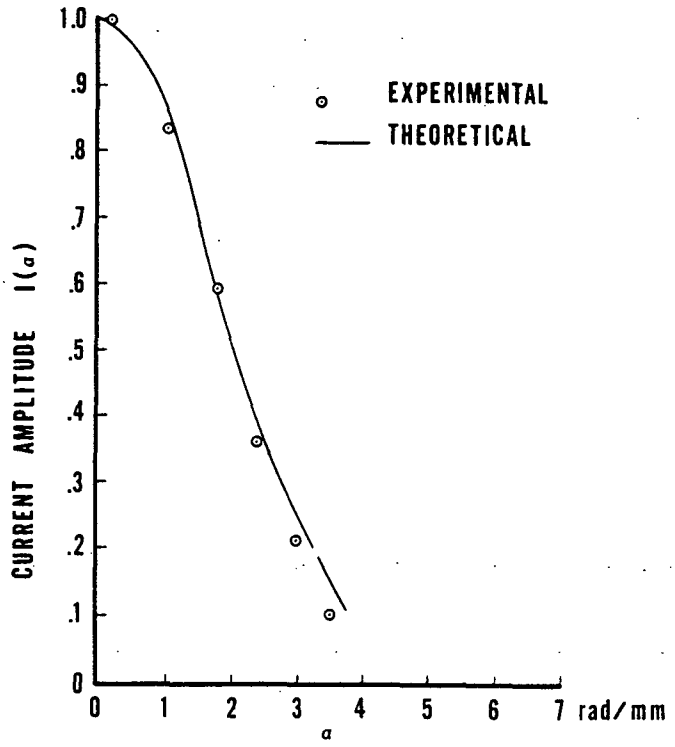


FIGURE 14. VALUES OF NORMALIZED HETERODYNE CURRENT VERSUS α FOR $W=\infty$, $a = 1.26\ mm^{-1}$

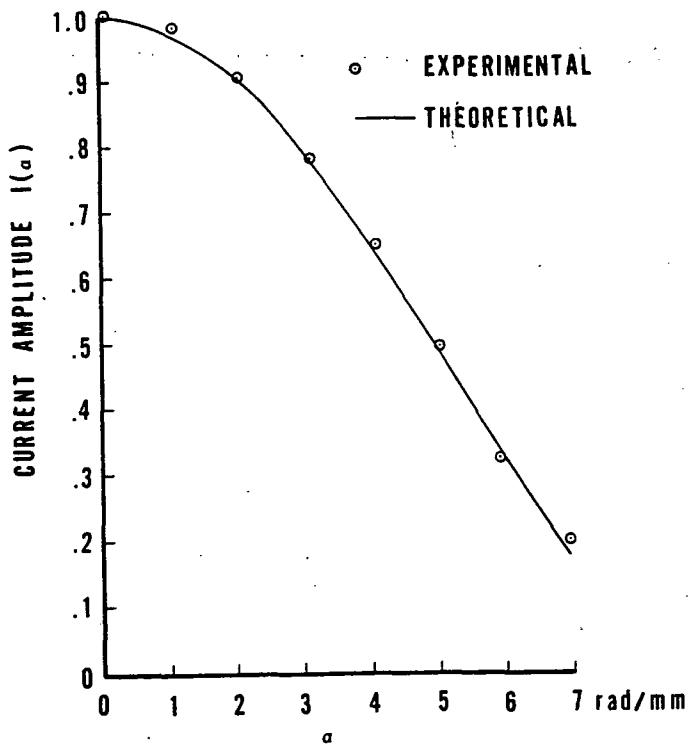


FIGURE 15. VALUES OF NORMALIZED HETERODYNE CURRENT VERSUS α FOR $W = 0.4\ mm$, $a = 1.26\ mm^{-1}$

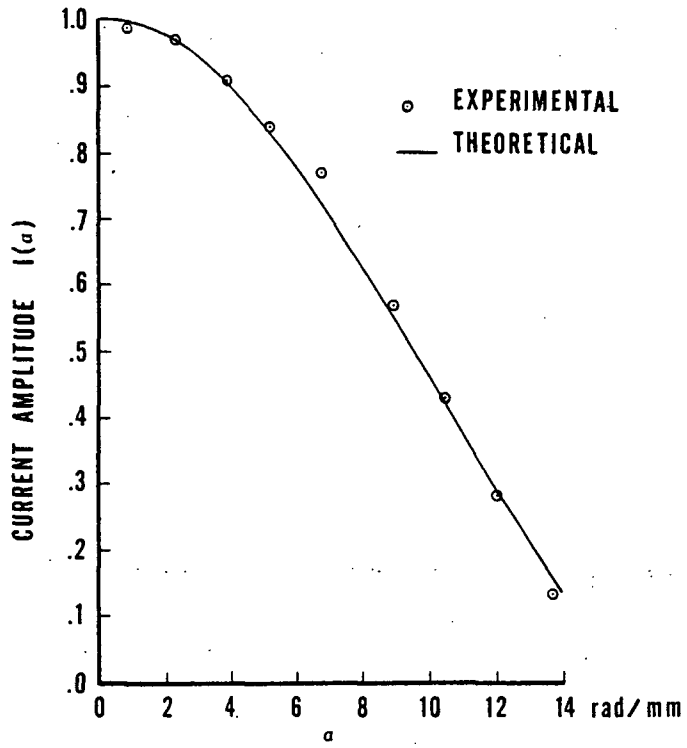


FIGURE 16. VALUES OF NORMALIZED HETERODYNE CURRENT VERSUS α FOR $W = 0.2\ mm$, $a = 1.26\ mm^{-1}$

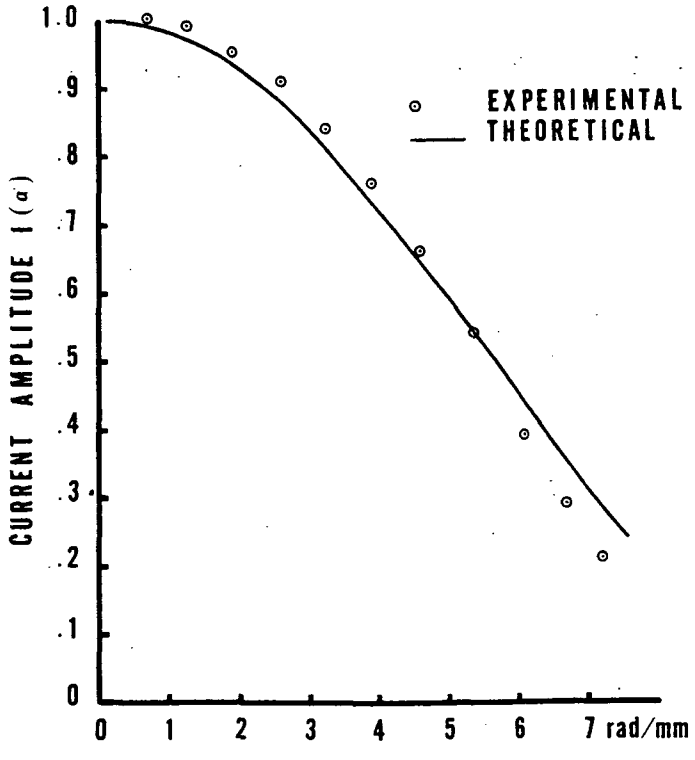


FIGURE 17. VALUES OF NORMALIZED HETERODYNE CURRENT VERSUS α FOR $R_0 = 0.4\ mm$, $a = 1.26\ mm^{-1}$

experimental values of normalized heterodyne current as a function of angle α for a circular aperture of radius R_o = 0.4 mm and a = 1.26. Generally, good agreement was obtained.

EXPERIMENTAL INVESTIGATION

A block diagram of the experimental setup used in this investigation is shown in Figure 18. A gas laser operating continuously at a wavelength of 6328Å was used as a coherent light source. The laser beam was passed through a collimator so that the beam width could be varied, thus varying the Gaussian field intensity distribution and also producing parallel wave fronts.

To obtain a difference frequency, a beam splitter was used to direct part of the output onto a fixed mirror and back to the cathode of a photomultiplier. The other part of the beam was directed onto a moving mirror and reflected back to the photomultiplier cathode. The moving mirror was translated back and forth sinusoidally. When the mirror was moving toward the light beam, the reflected beam had a frequency of $f_o + f_d$. This light beam was then mixed with the beam from the fixed mirror of frequency f_o to give a difference frequency of

$$f_d = 2v/\lambda_o.$$

This output signal was passed through a squaring circuit and then rectified to measure its amplitude. By rotating the fixed mirror through a given angle and measuring this angle with an autocollimator, a known angle between the two beams was induced. Rotating the fixed mirror through an angle caused a displacement on the photocathode. To compensate for this displacement and bring the two beams in coincidence on the photocathode, an optical flat was placed at an angle in the return path of the beam from the fixed mirror. By using rooftop reflectors for both the fixed and moving mirrors, it was possible to offset the return beam so the optical flat could be placed in the return beam path only.

The output of the gas laser is almost Gaussian when operating in the TEM_{oo} mode. This field intensity distribution was measured with an xy-recorder. A slit was placed in front of a photomultiplier and

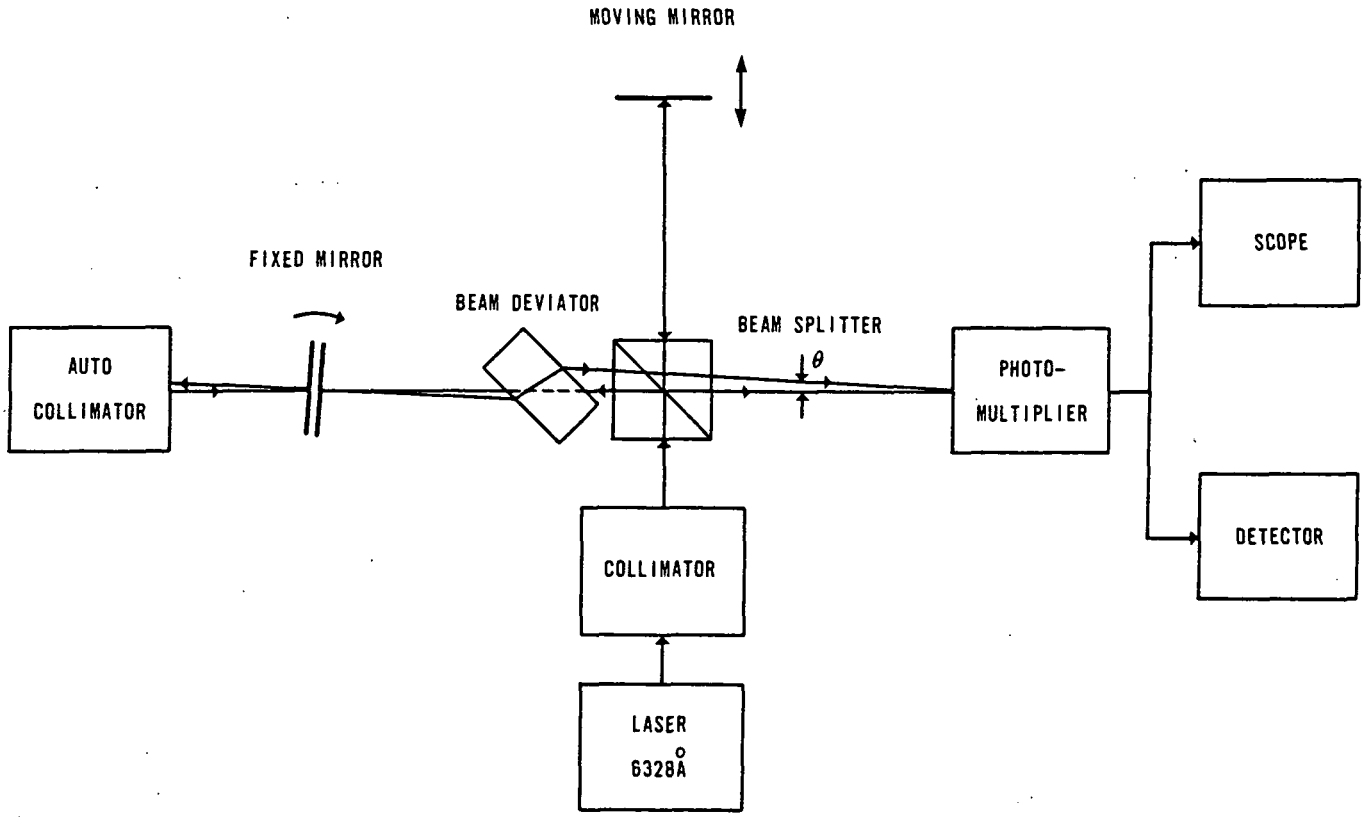


FIGURE 18. BLOCK DIAGRAM OF EXPERIMENTAL SETUP

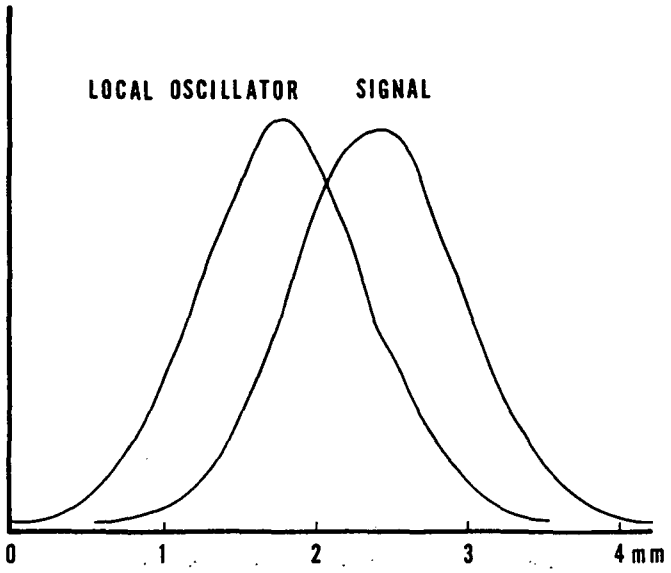


FIGURE 19. EXPERIMENTALLY DETERMINED FIELD INTENSITY DISTRIBUTION

driven across the laser beam by a motor which also drives a Helipot for the recorder x-axis. The photomultiplier output was connected to the y-axis of the recorder giving the intensity distribution as shown in Figure 19. The standard deviations for the local oscillator and the signal were then calculated from these curves.

CONCLUSIONS

Two coherent light sources of slightly different frequency may be mixed to give a beat frequency, provided the two beams are almost parallel to each other and overlap on the photocathode surface. It was found that the field intensity distribution of the light beams has a large effect on heterodyning. A Gaussian intensity distribution will give photomixing over a larger angle than a uniform field intensity distribution. When the two beams are not exactly parallel to each other, the detector aperture should be restricted to obtain maximum heterodyne current. By collimating or focusing the beams down to a small spot, most of the energy may be contained in the width $2w = \pi/\alpha$ mm for relatively large values of α. However, when focusing is used, care must be taken to place the detector in the focal plane so the wave fronts will be parallel; otherwise, the wave fronts will not be plane and the resulting current will be reduced.

REFERENCES

1. Boyd, G. D., and Gordon, J. P., "Confocal Multimode Resonator for Millimeter Through Optical Wavelength Masers," The Bell System Technical Journal, March 1961.

2. Fox, A. G., and Li, Tingye, "Resonant Modes in a Maser Interferometer," The Bell System Technical Journal, March 1961.

3. Boyd, G. D., and Kogelnik, "Generalized Confocal Resonator Theory," The Bell System Technical Journal, July 1962.

4. Born, Max, and Wolf, Emil, Principles of Optics, Pergamon Press, New York, 1959.

5. Sokolnikoff, Ivan S., Advanced Calculus, McGraw-Hill Book Company, Inc., New York, 1939.

"Page missing from available version"

"Page missing from available version"

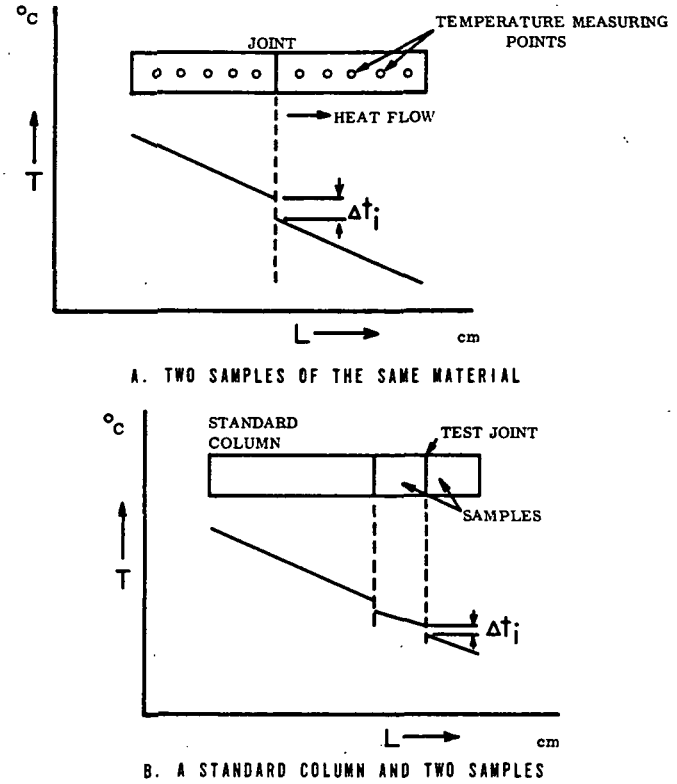


FIGURE 3. HEAT FLOW THROUGH TEST COLUMN

2. Repeatability from test to test is necessary to permit a reasonable comparison of data. With external radiation control, identical temperature differentials in the fluxmeter result in identical heat flow values.

3. Utilization of the overall testing concept where Δt_j is determined by calculation requires an accurate measurement of heat flow. The fluxmeter method of determining heat flow seems to be more accurate and desirable because it permits heat flow measurements to be made close to the test joint. For this purpose a second basic term, the conduction within a solid prismatic body, must be introduced.

$$Q_s = \frac{k \cdot A \cdot \Delta t_s}{L} \qquad (2)$$

where

Q_s = heat flow normal to the cross sectional area (watts)

k = conductivity of the solid $\frac{\text{watts}}{\text{cm}^\circ\text{K}}$

Δt_s = temperature differential between two points (°K)

L = distance between these points (cm).

Basically the mating solids can serve as the "standard" column for determining Q_s (Fig. 3A); however, a great variety of materials with varying and poorly established conductivity values was to be tested, so a standard column was introduced as shown in Figure 3B.

This column was a piece of Armco iron, chosen because it is a highly pure material with a relatively low conductivity which provides a high temperature gradient. A question was raised about the accuracy of the Armco iron conductivity. Even though published data were used in the calculations, conductivity of the Armco iron will be verified by test. Both samples and the standard (in the following called fluxmeter) have the same circular cross sectional area. If no lateral heat losses occur, the flow Q_j must be equal to the flow Q_s.

$$Q_s = Q_j. \qquad (3)$$

Substituting Q_s for Q_j in equation 1b yields

$$h = \frac{Q_s}{\Delta t_j \cdot A} \quad \frac{\text{watts}}{\text{cm}^2 \,^\circ\text{K}}. \qquad (4)$$

This equation presents the basis for constructing the test apparatus.

TEST APPARATUS

The three main parts of the test apparatus are the test fixture, instrumentation, and vacuum system (Fig. 4).

A. Test Fixture

The test fixture shown in Figure 5 was constructed of two flat plates held together by three steel rods, a pressure bellows mounted to the bottom plate, and an adjusting screw in the top plate. The use of a bellows was recommended by E. Fried (Ref. 1) for convenient changing of the contact load without disturbing the vacuum. Going from bottom to top (in the direction of the heat flow), the central or actual measuring column consisted of the main heater, fluxmeter, lower sample, upper sample, cooler, and load cell.

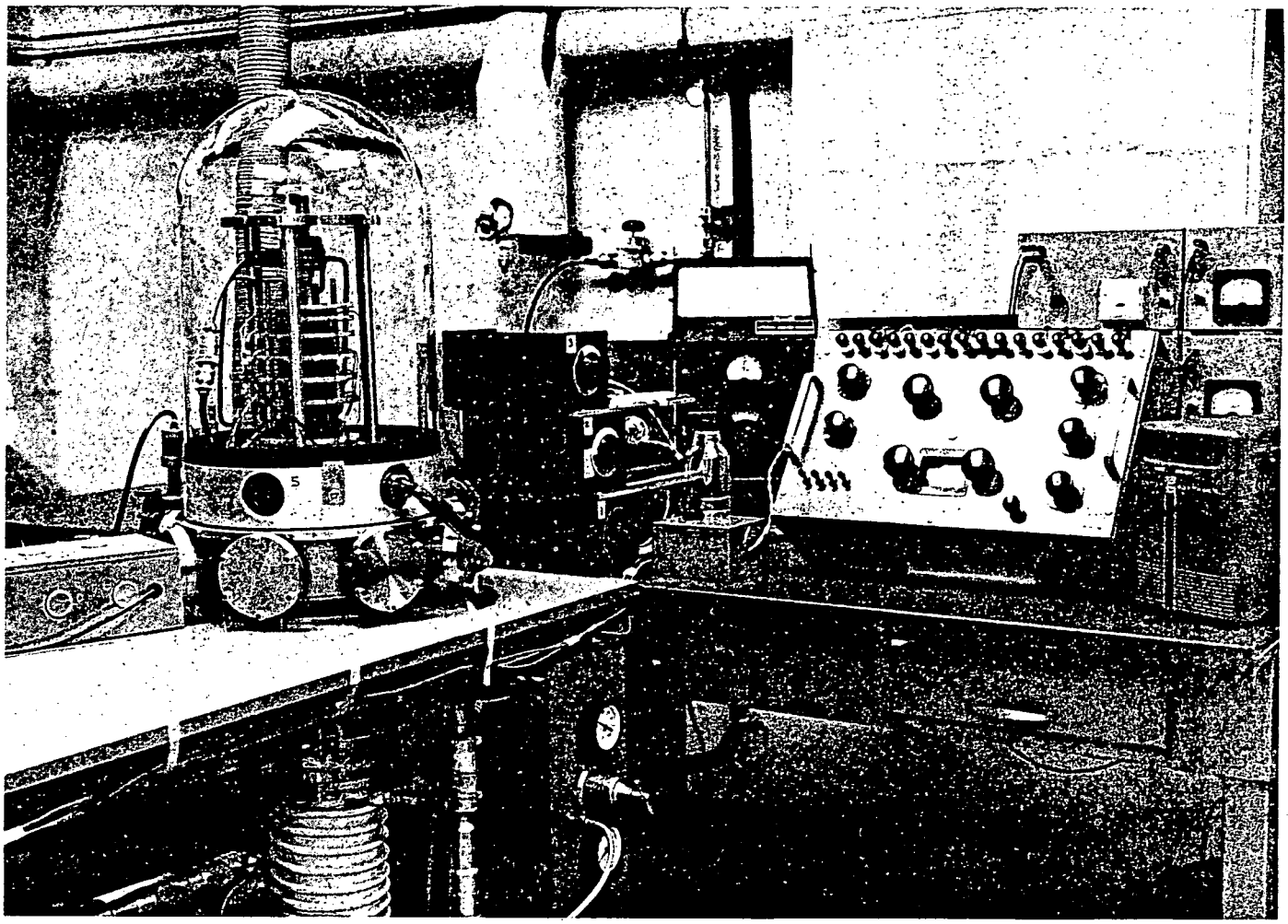


FIGURE 4. TEST APPARATUS

All contacts where a low heat resistance was desired were lapped and filled with high vacuum silicone grease; these were the joints between (1) heater and fluxmeter, (2) fluxmeter and lower sample, and (3) upper sample and cooler.

To reduce the lateral heat transfer from or to the center column by radiation, an improved shielding device was provided. The tendency has been for other experimenters to use a single shield with an encompassing heater. This limits the vertical temperature matching or nulling characteristics of the radiation shield.

The radiation shielding device is unique because of its ability to reduce the radiation losses in six discrete levels of the test column. Gold-plated aluminum rings bonded with graphite-cloth heating elements were each controllable so that mating the temperatures of the rings with those of the column would reduce radiation to an insignificant minimum.

It was learned from previous tests that the rings influence each other by a "mirror" effect on the column; therefore, separating disks of fiberglass were introduced to serve as radiation heat baffles. The diameter of the cylindrical column was 3.8 cm, the length of the fluxmeter was 8.2 cm, and the length of the samples was 1.9 cm.

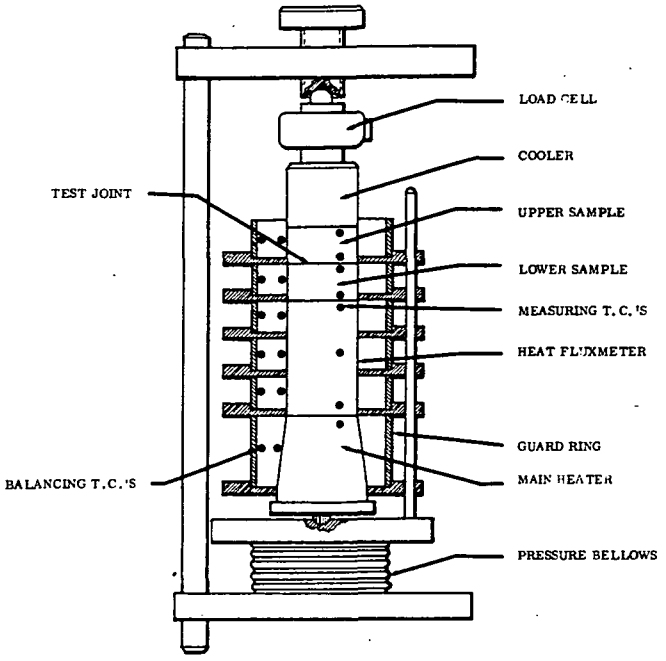


FIGURE 5. TEST FIXTURE

B. Instrumentation

As previously shown, temperature measurements throughout the test fixture were very critical. The problems associated with accurate temperature measurement became apparent as the test apparatus design progressed.

Figures 5 and 6 give some insight as to the complexity of the thermocouple instrumentation. Several of the more important factors and the basic reason given for their careful consideration during instrumentation of the test fixture are described.

1. Thermocouple instrumentation was considered the most critical factor in the test fixture design. The calculation of Q_S (equation 2) utilizes a Δt_s measured at two specific levels as shown in Figure 5. In addition to providing the basis for determining h with the calculated heat flux, those thermocouples located concentrically around the test column were utilized to assist in monitoring and reducing lateral heat losses.

2. The lateral heat losses were reduced to a negligible quantity by the use of the guard rings (Fig. 5). The temperatures of the guard rings were continuously monitored during tests and adjusted to insure that the temperature difference between the test column surface and the inner surface of the guard rings was held to less than 1° K (equivalent to 40 microvolts). Controlling the temperature differential to 1° K or less will reduce the radiation losses from the test column to a value of approximately 0.002 watts or 0.007 percent of the input electrical power. This is a negligible quantity since the effect upon the calculated conductance will be minimal. To further insure compatible temperature readings, thermocouples on the test column and the guard ring surfaces were cemented in place; i.e., identical mounting techniques were used.

Monitoring and adjusting the guard rings will be eliminated when the automatic guard ring controller is fabricated. An electronic circuit has been designed to compare the outputs of two heat sensors (one on the test column and one on the guard ring) and then to adjust the power input to the guard ring to null the readings to less than 1° K difference.

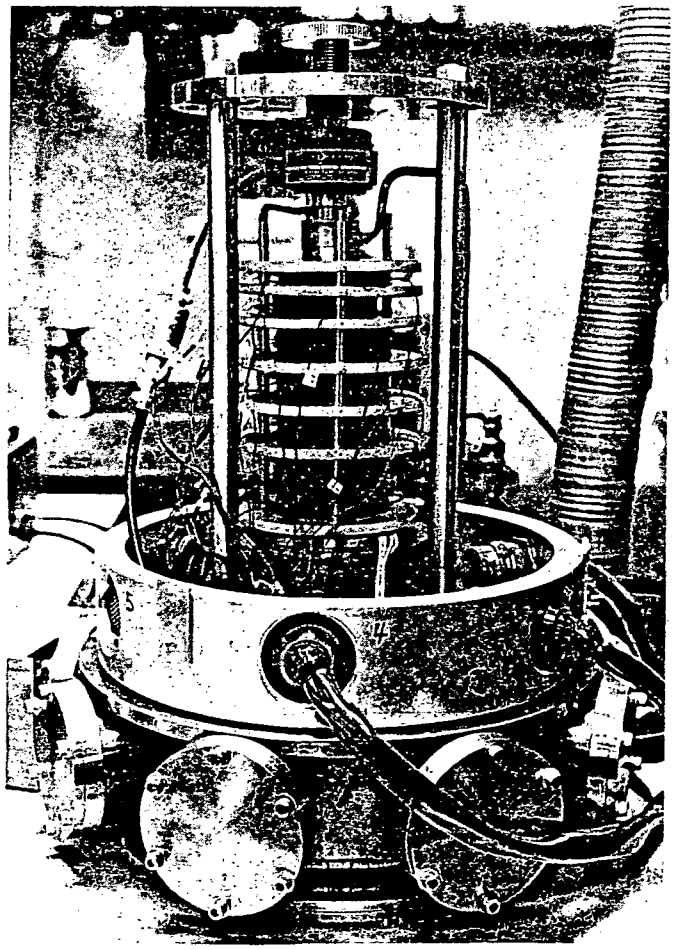


FIGURE 6. TEST FIXTURE

3. The initial thermocouples used in the test program were calibrated and found to deviate less than 0.1°K from NBS standards. All thermocouples to date have been from the same manufacturer and, to the best of the authors' knowledge, from the same material batch; therefore, their accuracy is considered adequate. All reasonable precautions have been taken to insure temperature measurement accuracy.

4. The temperature differential (Δt_j) across the joint, i.e., between the mating surfaces of the test samples, was determined in the following manner.

Thermocouples were inserted in the samples in the pattern shown in Figure 7, at an average distance of 0.16 cm from the test surface. This average was obtained by X-ray and by measurement under a microscope.

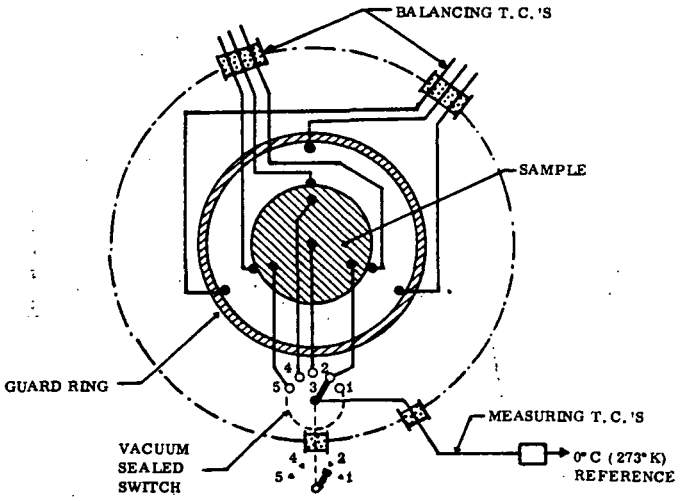


FIGURE 7. ARRANGEMENT OF THERMOCOUPLES

A temperature differential from the thermocouple to the test surface could be computed with the previously calculated heat flow (watts), the measured distance (L) from the thermocouple to the surface, the known cross sectional area, and the published conductivity of the test sample. This value was then either subtracted from the measured temperature for the lower sample or added for the upper sample to obtain the surface temperatures.

A question was raised concerning the placement of thermocouples so close to the sample test surface. Reference 2 indicated that the thermocouples could be within the area of constriction resistance where the heat flow lines in the material are being directly affected by the metallic contacts.

To verify the procedure described, two samples as shown in Figure 3A will be prepared. After completion of the test, temperature measurements will be plotted as shown in Figure 3A, and the Δt_j obtained from this plot will be used to calculate h. The variation between this h and that calculated from using the short samples will be indicative of the error introduced by thermocouple location.

For ease in fixture maintenance and changing of samples, a rotary vacuum feed-through switch was utilized (Fig. 7). Thermocouple wires from the column were soldered to the switch points inside the vacuum chamber. A single thermocouple was used to reference the switch to an ice-bath junction outside the vacuum jar. This technique permitted a complete thermocouple circuit to exist between the measuring point and the potentiometer.

EXPERIMENTAL PROGRAM

An experimental program was set up according to the design concept of cold plates to be used in Saturn application (Table I).

At the inception of the testing program, the prime material under consideration for the cold plate mounting surface was Aluminum 6061 T6; therefore, the colder (upper) samples were made of this material in an attempt to simulate cold plate operation.

TABLE I. THERMAL CONTACT CONDUCTANCE TESTING PROGRAM

Run No.	Lower Sample		Upper Sample		Remarks
	Material	Finish*	Material	Finish	
1	Almag 35	Lapped	6061-T6	Lapped	
2	Almag 35	0.81-1.60	6061-T6	0.81-1.60	
3	Almag 35	2.54-5.08	6061-T6	2.54-5.08	
4	Almag 35	2.54-5.08	6061-T6	2.54-5.08	
5**	Mag AZ91C	0.81-1.60	6061-T6	0.81-1.60	Hysteresis testing
6	Almag 35	2.54-5.08	6061-T6	2.54-5.08	Silicone grease interstitial layer
7	Mag AZ91C	0.81-1.60	6061-T6	0.81-1.60	
8	Alu 356	0.81-1.60	6061-T6	0.81-1.60	
9	Mag AZ91C	Black Anod. 0.81-1.60	6061-T6	0.81-1.60	
10	Almag 35	0.81-1.60	Mag AZ91C	0.81-1.60	

* The finish refers to microns CLA machined with flycutter except for run 1.

**Run 5 increasing and decreasing contact pressure until variation in conductance is negligible.

Cast alloys such as Almag 35, Magnesium AZ91C, and Aluminum 356 were chosen for the hot (lower) samples because of their extensive use in the Saturn instrument unit.

Table I shows test run 6 with an interstitial layer. More elaborate testing with interstitial layers will be carried on as shown in Table II. Also the behavior of soft foils, such as indium, will be studied separately (run 14).

TABLE II. THERMAL CONTACT CONDUCTANCE TESTING PROGRAM

Run No.	Lower Sample		Upper Sample		Remarks
	Material	Finish*	Material	Finish	
11	Alu 356	Fly** 0.81-1.60	Mag AZ91C	0.81-1.60	
12	Almag 35	Endmill 0.81-1.60	6061-T6	Endmill 0.81-1.60	
13	Almag 35	Shaper 0.81-1.60	6061-T6	Shaper 0.81-1.60	
14	Almag 35	Fly 0.81-1.60	6061-T6	Fly 0.81-1.60	Foil investigation indium
15	Almag 35	Lapped Black Anod.	6061-T6	Lapped Black Anod.	
16	Almag 35	Lapped	Mag AZ91C	Lapped	Hysteresis test
17	Al2024T3	Lapped	Mag AZ91C	Lapped	
18	Almag 35	Shaper 0.81-1.60	Mag AZ91C	Fly 0.81-1.60	
19	Almag 35	Endmill 0.81-1.60	6061-T6	Fly 0.81-1.60	Foil investigation aluminum
20	Alu 356	Fly 0.81-1.60	Mag AZ91C	Fly 0.81-1.60	Silicone grease interstitial layer

* Finish as indicated by process and micron CLA.

** Fly is flycutter.

TEST RESULTS

Test run 1 was performed mainly to estimate the accuracy of the results which could be expected with the described apparatus. To obtain a good joint, the test surfaces of the samples were carefully lapped.

No temperature difference should occur between any two measuring points within one horizontal plane in the case of the ideal conditions, i.e., no lateral heatflow, homogeneous materials within the test column, and a perfect joint.

The lateral heatflow could be kept at an insignificant amount by maintaining the guard ring temperatures as previously mentioned. The homogeneity of the materials involved was assumed to be sufficient.

Each of the thermocouples designated as measuring T.C.'s in Figure 7 was monitored during the test run. The temperature readings within each of the measuring planes differed not more than 0.5° C. The two regions of the column used for conversion of the temperature readings to conductance values were (1) the center region, represented by the single thermocouples in the center, and (2) the outside region, represented by the average temperatures of the three outer thermocouples (Fig. 7). The two corresponding conductance values were in good agreement. The difference between the center and outside conductance was approximately 10 percent at lower contact loads and approximately 5 percent at higher loads.

The contact conductance versus contact pressure plotted in Figures 8, 9, and 10 are based on the average of center and outside values.

The data were obtained from samples which had not been mated until they were assembled in the test fixture. Samples were machined, wrapped, and stored for periods up to four months. Oxide coatings that

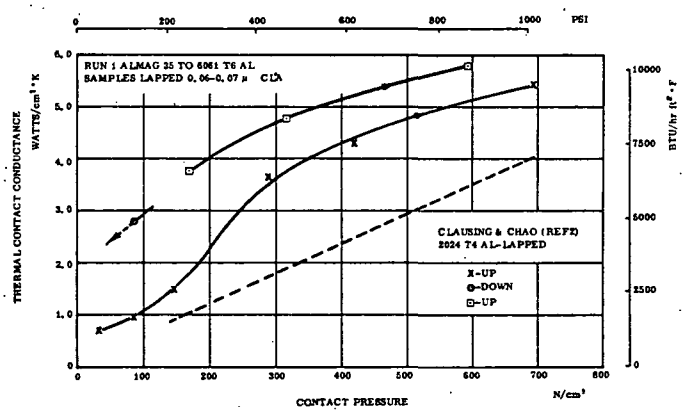


FIGURE 8. CONDUCTANCE VERSUS PRESSURE, RUN 1

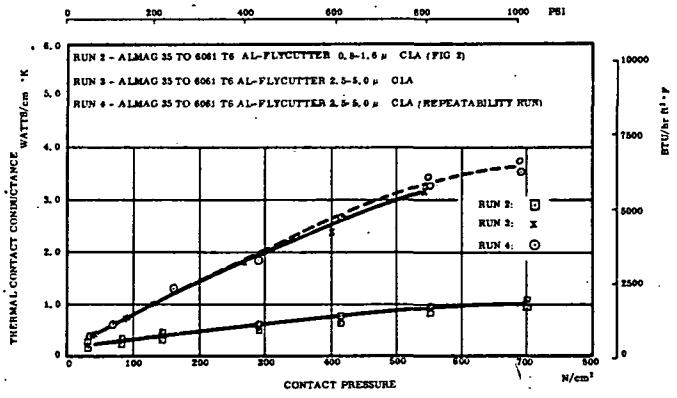

RUN 2 - ALMAG 35 TO 6061 T6 AL-FLYCUTTER 0.8-1.6 μ CLA (FIG 2)
RUN 3 - ALMAG 35 TO 6061 T6 AL-FLYCUTTER 2.5-5.0 μ CLA
RUN 4 - ALMAG 35 TO 6061 T6 AL-FLYCUTTER 2.5-5.0 μ CLA (REPEATABILITY RUN)
RUN 2: ⊡
RUN 3: X
RUN 4: ⊙
THERMAL CONTACT CONDUCTANCE
WATTS/cm²·K
BTU/hr ft²·F
CONTACT PRESSURE
N/cm²
PSI

FIGURE 9. CONDUCTANCE VERSUS PRESSURE, RUNS 2, 3, AND 4

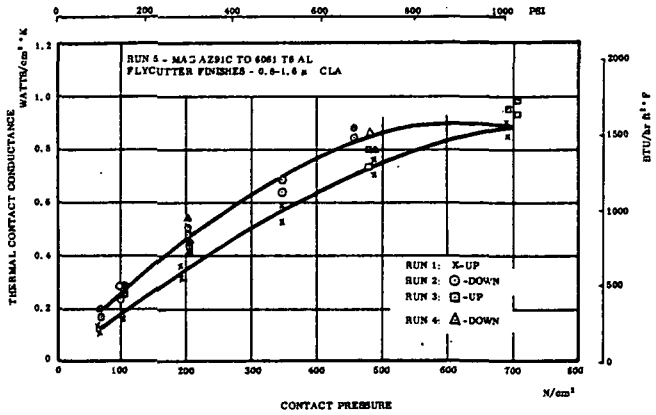


FIGURE 10. CONDUCTANCE VERSUS PRESSURE, RUN 5

may have formed during this time period were not disturbed. This was done to simulate storage of Saturn components.

Results from test run 1 with hysteresis points (Fig. 8) show that the conductance values are greater than those in Reference 2. This difference can possibly be traced to difference in materials, surface finish measurement techniques, and even to the testing methods themselves. These items indicate the difficulty to be experienced in comparing data from different experimenters.

It is generally accepted that the finer the surface finish, the greater the contact conductance. This assumes the roughness as the sole parameter of comparison.

The data from test runs 2, 3, and 4 (Fig. 9), however, ran contrary to this statement. At a contact pressure of 550 N/cm^2, for instance, the coarse-finish joint (run 2) displays a conductance of 3.3 watts/cm^2 °K, while the fine-finish joint (runs 3 and 4) shows only 0.9 watts/cm^2 °K.

An inspection of the test surfaces disclosed a considerable surface waviness of the samples in run 2 (Fig. 2). The maximal wave height (peak to valley distance) was approximately 20 μ. The wave height of the surfaces in the high-conductive joint, however, was only 3 μ. This value practically disappears within the finish roughness of 2.5 to 5 μ.

A comparison of the surface conditions and the conductance results from runs 2 and 3 shows that the waviness (or flatness deviation) of the joint faces has a great effect on contact conductance. For obtaining conclusive results, the description of a surface must give both roughness and waviness.

However, sufficient experimentation has not been done to determine the magnitude of the effect of surface flatness on contact conductance. Profile measurements taken after the application of contact pressures exceeding 700 N/cm^2 did not disclose any significant changes in finish roughness and flatness deviation. The appearance of certain regions on the test surfaces after the test could not be readily explained.

Runs 3 and 4 utilized identical materials, finishing processes, and relative flatness. These runs were made to verify repeatability characteristics of the test apparatus. As indicated in Figure 9, less than 5 percent variation was found between runs 3 and 4.

Run 5 (Fig. 10) was performed primarily to determine the change in contact conductance with repeated cycling of pressure. Mounting bolt area characteristics are such that the macroscopic contact will remain the same; however, the microscopic contact condition cannot be repeated by removing and subsequent remounting to the same location. In this case, the conductance will not vary from the initial value as indicated by the lower curve in Figure 10.

CONCLUSIONS

Since the primary purpose of this experimental program was to determine design data for Saturn vehicle components, review of microscopic effects upon contact conductance is suggested for a future program.

Test data shown in Figures 9 and 10 indicate that the magnitude of thermal contact conductance attainable with normal machined surfaces in the Saturn program will not create excessive temperature differentials between the conditioning panel and the component mounting surface. For design purposes, a heat load of 10 watts per mounting bolt area (5 × 5 cm) has been assumed. Under this condition, utilizing the contact conductance from run 2 at 350 N/cm^2 of 0.7 watts/cm^2 °K, a temperature differential of less than 1° K would exist. For component thermal design purposes, this differential can be neglected in most cases.

Because of discrepancies, run 2 data cannot be used for design purposes. A separate program will be necessary to determine the relationship of flatness deviation to contact conductance. A program of this

nature will require a review of the methods used for measuring temperature differentials across the test joint.

Data obtained from the present experimental program cannot be utilized to determine an average temperature differential over the entire component mounting surface because of the lack of knowledge concerning the average contact pressure. Presumably, the only definable contact pressure is directly adjacent to the mounting bolt.

Prior to runs 2 and 3, profilometer readings were taken to obtain some insight into the macroscopic and microscopic characteristics of the test surfaces. Measurements made after a test cycle did not indicate major changes in the surface characteristics.

It was concluded that surface changes could not be evaluated with profilometer measurements. This conclusion is supported by the work of Fenech and Rohsenow in Reference 4.

Test runs to date (1 through 5) and the remainder scheduled in Tables I and II will not permit full understanding of the effect of the microscopic character of the surface on thermal contact conductance. Many parameters cannot be properly evaluated with the relatively few test runs scheduled.

A greatly expanded program coordinated with programs of other experimenters is necessary to fully evaluate the mechanics of thermal contact conductance in a vacuum.

The following items are considered the major areas of interest to be studied in a future program:

1. The relationship between surface characteristics (i.e., flatness deviation and surface finish) and contact conductance. To properly compare test data, a common basis for defining surface characteristics must be agreed upon. To evaluate flatness deviation and surface finish, each must be taken separately and then combined in a closely controlled test.

2. Investigation of the problems involving the standardization of test samples, i.e., diameter, length, surface measurement technique, and material thermal conductivity. A single source to supply samples would be a good solution to the problem of discrepancies in materials and methods of surface measurements.

3. Temperature measurement techniques and data evaluation procedures should be standardized to the greatest extent practicable.

4. The effect of nonmetallic interstitial materials on the conductance should be thoroughly investigated.

5. Conductance variation when mating similar materials and dissimilar materials must be determined.

REFERENCES

1. Fried, Erwin, and Costello, Frederick A., "Interface Thermal Contact Resistance Problem in Space Vehicles," ARS Journal, February 1962, pp. 237-243.

2. Clausing, A. M., and Chao, B. T., "Thermal Contact Resistance in a Vacuum Environment," University of Illinois, Engineering Experiment Station, Report ME-TN-242-1 NASA Research Grant NsG-242-62.

3. Prepared under sponsorship of National Aeronautics and Space Administration, United States Air Force, and United States Weather Bureau, "U. S. Standard Atmosphere, 1962," U. S. Government Printing Office, Washington, D. C., December 1962, p. 13.

4. Fenech, H., and Rohsenow, W. M., "Thermal Conductance of Metallic Surfaces in Contact," Massachusetts Institute of Technology, U. S. Atomic Energy Commission, Contract AT(30-1)2079, May 1959.

D. FLAT CONDUCTOR CABLES by Wilhelm Angele

Cables made of flat conductors are compared with those made of round wires. Weight, cost, reliability, and time required for assembly are discussed. Need for further development of connectors and of handtools for installing connectors is noted.

INTRODUCTION

In the flat conductor cable, the conductors are flat ribbons of copper laid parallel in the same plane and encased in a dielectric plastic material (Figs. 1 and 2). Flat conductor cables of various configurations are commercially available. Materials such as polyesters, fluorocarbons, or H film are generally used for the plastic insulation film and determine the operating temperature range of the cables (-65°C to 200°C). The high dielectric and tensile strength of these materials are the principal reasons for the superiority of flat conductor cables over cables of round wire with polyvinyl chloride (PVC) insulation. The use of these materials permits much thinner insulation, which contributes up to 32 percent weight saving. Molding the cable termination instead of soldering and potting can reduce the cost of a cable up to 80 percent. Installation time is much less than for round cables because contact adhesives are used for mounting instead of clamps, brackets, etc. Reliability is improved by the use of machine tools rather than handwork and also by elimination of solder joints. Specifications and standards for cables and connectors have been established and hardware has been successfully tested. The growing interest in flat conductor cables necessitates this comparison of round and flat conductor cables. Flat conductor cables are being developed to keep pace with electronic parts progress, circuitry design, and modern packaging techniques.

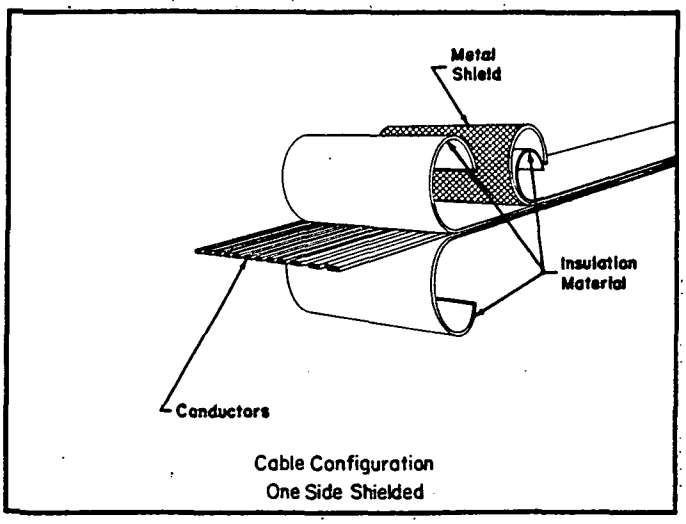


FIGURE 1. SHIELDED FLAT CONDUCTOR CABLE

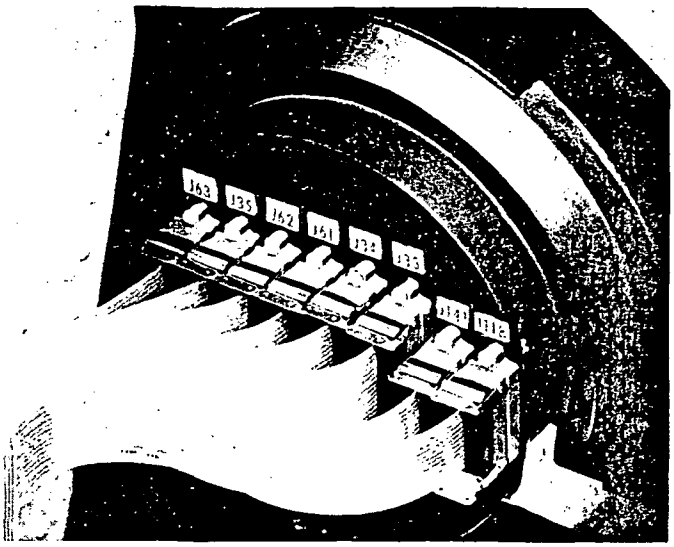


FIGURE 2. FLAT CONDUCTOR CABLES INSTALLED

Cables are produced with various insulation materials, conductor sizes, numbers of conductors, and conductor spacings. They may also be produced with shields of various materials and configurations. Electrostatic shields made of metals or conductive paint can be applied by foil or wire screen lamination, vacuum deposition, or painting. The shield may be solid, perforated, wrinkled, or cut and expanded. The shields may be applied at various distances from the conductors on one or both sides of the cables.

INSULATION MATERIALS

One main reason for the flat conductor cable's superiority over round wire cables is the excellent insulation material. Because of their high tensile and dielectric strength, Mylar and H film allow thinner insulation than PVC, thus contributing to mass and space savings.

Mylar can be used up to 100°C continuously. The temperature rating of the Teflon-FEP-bonded H film

cable is limited by the Teflon to 200°C. Addition of fiberglass to the Teflon adhesive raises the temperature rating to 250°C. The low dielectric constant of the Teflon is desirable. A new adhesive system is needed whose temperature rating is equal to that of the H film (600°C).

TABLE I. INSULATION MATERIALS

Properties	Unit	Mylar	Teflon-FEP fiberglass	H film	PVC
Dielectric strength	10^6 V/cm	2.4	.4	2.4	0.3
Dielectric constant (10^3 Hz)	-----	3.2	2.5 to 5*	3.2 to 3.5	3 to 4
Tensile strength	10^6 N/m^2	124	40 to 100*	120	20
Maximum service temperature	°C	125	250	500	85
Gamma radiation endurance	REP	10^8	10^5	5×10^9	10^7

*Depending on the amount of fiberglass.

BASIC CABLE LAYOUT AND DESIGN CONSIDERATIONS

Cables should be designed in advance to obtain full benefit of the flat conductor cable's simplicity. Straight runs with one connector at each end are recommended. A connector may accommodate more than one cable termination. Branching of cables should be used when single straight cables are not advisable.

Another important design consideration is the pin sequence in connectors of corresponding boxes. Naturally, this can be coordinated with little effort when designing new boxes. Changes of pin sequence can be made without entering the box by using a special wiring plug at one end of the cable.

Boxes equipped with round wire standard connectors may be interconnected with flat cables by slitting, soldering, and potting the cable ends into conventional plugs.

CABLE END PREPARATION

Tools for stripping the flat cable ends preparatory to attaching the plug have been built by several cable and connector manufacturers. Generally, the two basic approaches used are the skiving blade and chemical etching. A semiautomatic machine that strips one cable end in only 30 seconds has been built. The plastic (polyester and some fluorocarbons) is shaved off by a heated blade which is continually cleaned by a paper strip. The plastic remaining between the conductors is removed by a second paper strip. Flat cables with H film insulation and Teflon bond are stripped without heat in a few seconds simply by using a handtool.

The stripping of shielded cable is more complex. A simple, practical method has not been developed, although with care the heat method described has been used successfully.

After the cable end has been stripped, the plug is conveniently formed by injection molding. The total time to prepare a cable end and to mold a plug is 2 to 3 minutes. The time required to strip, solder, test, and pot a plug of the type used in conventional round wiring is 2 to 3 hours; to wire a socket also requires hours. The 2 to 3 hours for attaching the plug does not include the harness work on the nailboard nor the lacing which is again a matter of hours. No additional time is required to prepare a printed circuit board to be plugged into a flat conductor cable connector. A flat conductor cable is cut and assembled in minutes.

INSTALLATION MATERIAL, TECHNIQUE, AND PROCEDURE

The flat conductor cable can most conveniently be installed with adhesives rather than by clamps and brackets. Several adhesive systems such as thermosetting, thermoplastic, solvent activated, and contact adhesives have been considered and tested. The design requirements specify certain peel and creep strengths through a large temperature range, high and low humidity, vibration, and many other environmental conditions. It appears, at present, that a pressure-sensitive polyester adhesive is most practical. A liquid adhesive such as a nitrophenolic can also be used. A contract is underway to further develop a system applicable to Mylar, Teflon, and H film.

The routing path should be marked, and the cable should be numbered to facilitate processing, stocking, and installation. All sharp corners, edges, or other hazards in the path of the cable should be rounded, covered, or removed to prevent damage. The cable may be installed in layers. A number of cable fold patterns have been established to change polarity from right to left, to change the layer sequence of the cables, or to route the cable in any desired direction, including rotation about the cable's longitudinal axis (Fig. 3).

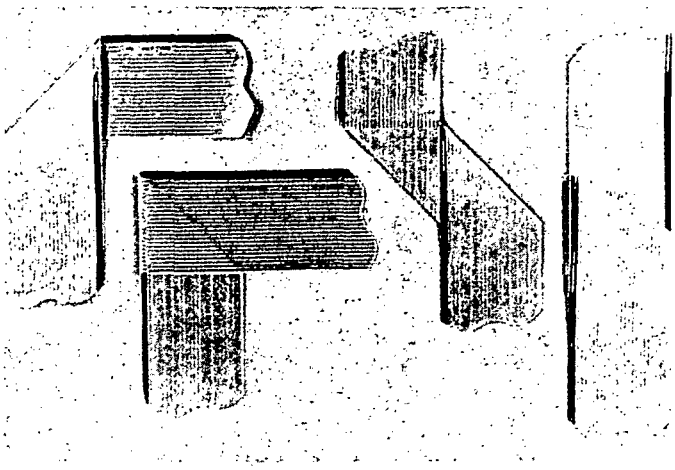


FIGURE 3. FOLDING TECHNIQUES

FLAT CONDUCTOR CABLE CONNECTORS

The design philosophy of flat conductor cable connectors is quite different from that used with most pin-and-socket round wire connectors (Figs. 4 and 5). Emphasis is concentrated on contact force, specific contact pressure, contact form, spring constant, tolerances, and metals. Contact resistance below one milliohm; long wear life; and proper functioning under shock, vibration, temperature extremes, and high humidity are required. Insertion and withdrawal forces of the plug are kept low because these forces

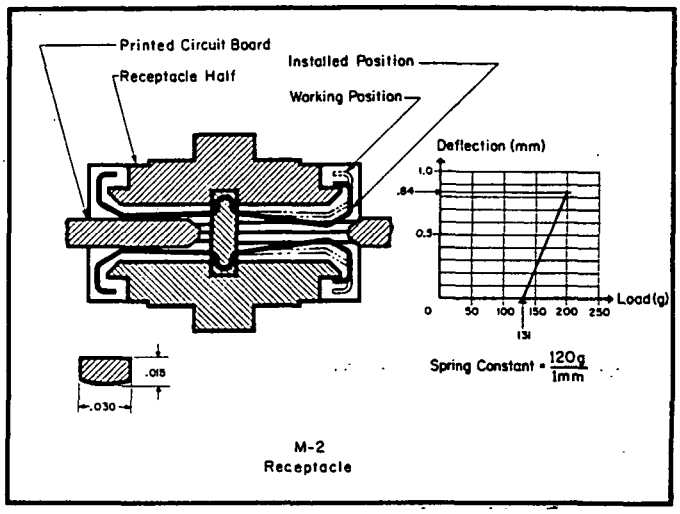


FIGURE 5. DOUBLE FEMALE CONNECTOR

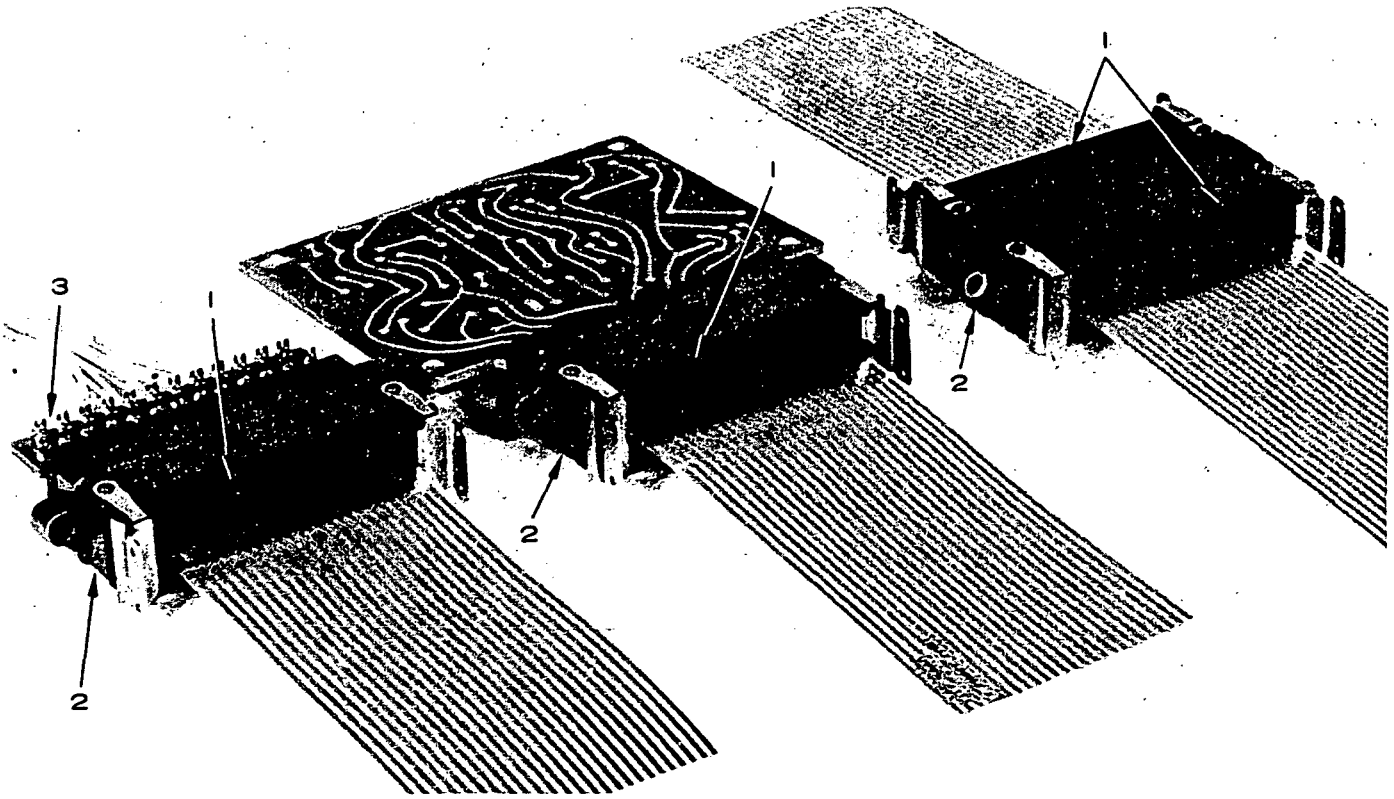


FIGURE 4. THREE TYPES OF CONNECTORS: (1) MALE CONNECTOR, (2) DOUBLE FEMALE CONNECTOR, AND (3) ROUND-WIRE ADAPTER

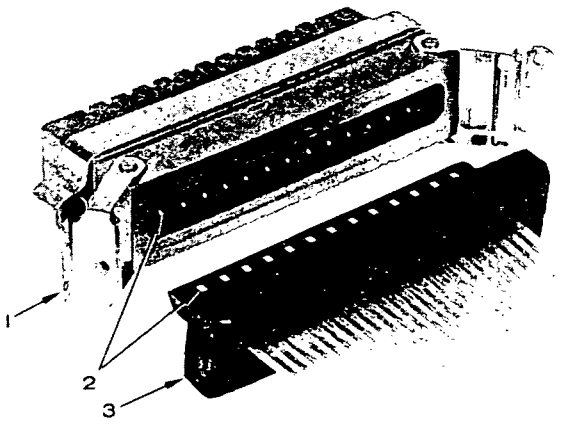


FIGURE 6. TYPICAL CONNECTORS: (1) SAFETY LATCH, (2) CONTACTS, AND (3) SEAL

are not considered important for the quality of a connection (Fig. 6). Polarity keys prevent wrong insertion, safety latches allow wiring against accidental opening, and seals protect the connector against moisture.

Table II compares the mass of standard round wire connectors with that of NASA flat cable connectors.

TABLE II. CONNECTOR MASS COMPARISON

	Standard Wire Connector 26 Pins	NASA 25 Pins	Saving
Mass per pin	2.84 g	1.56 g	45%
	Standard Wire Connector 61 Pins	NASA 50 Pins	Saving
Mass per pin	3.2 g	0.8 g	75%

Table III compares the connector mounting area required for standard round wire cables and NASA flat cables.

TABLE III. CONNECTOR-MOUNTING AREA

	Standard Round Wire Connector 26 Pins	NASA 25 Pins	Saving
Area per pin cm^2	.56	.47	16%
	Standard Round Wire Connector 61 Pins	NASA 50 Pins	
Area per pin cm^2	0.4	.236	41%

MASS COMPARISON

A. UNSHIELDED CABLES

The mass comparison is made for one conductor by using flat and round wire having similar copper cross section and electrical resistance as shown in Table IV. The average mass saving by using flat conductor cable in the wire range of AWG No. 20 to No. 32 is 25 percent. This does not include the additional saving obtained by replacing fastening hardware with adhesives.

TABLE IV. MASS CHART FOR INSULATED WIRE

Flat Wire Mylar Insulation						Round Wire (Stranded) Mil-W-16878 PVC Insulation		
Cond. Spacing c/c 10^{-3}in	Cable Thickness 10^{-3}in	Copper Dimen. 10^{-3}in	Copper Area $10^{-6}in^2$	Ohm/m	Mass g/m	AWG #	Area $10^{-6}in^2$	Mass g/m (average)
50	6	2 x 25	50	.547	.51	32	50	.82
	7	3	75	.365	.67	30	79	1.05
	8	4	100	.270	.83	29	100	1.26
	9	5	125	.218	1.01	28	126	1.53
75	6	2 x 25	50	.547	0.64	32	50	.82
	7	3	75	.365	0.83	30	79	1.05
	8	4	100	.270	1.01	29	100	1.26
	9	5	125	.218	1.21	28	126	1.53
75	6	2 x 40	80	.337	0.77	30	79	1.05
	7	3	120	.227	1.04	28	126	1.53
	8	4	160	.151	1.29	27	158	1.78
	9	5	200	.136	1.54	26	200	2.10
100	6	2 x 65	130	.208	1.17	28	126	1.53
	7	3	195	.139	1.57	26	200	2.10
	8	4	260	.104	1.97	25	252	2.53
	9	5	325	.084	2.38	24	320	3.10
	10	6	390	.070	2.78	23	400	3.66
150	6	2 x 115	230	.118	1.89	25	252	2.53
	7	3	345	.079	2.52	24	320	3.10
	8	4	460	.059	3.25	22	510	4.38
	9	5	575	.047	3.96	21	640	5.44
	10	6	690	.039	4.63	21	640	5.44
	11	7	805	.034	5.35	20	800	6.79
	12	8	920	.030	6.05	19	1000	8.43

NOTES: AWG up to #24 are considered standard for cabling.
AWG above #24 to #28 are limited to short runs on protected areas (inside of boxes).
AWG above #28 (fine wires) are not recommended for handmade cable harnesses.

B. SHIELDED CABLES

The shielding data of round and flat cables are cumbersome to compare because of the multitude of shielding configurations. Table V compares an insulated flat conductor having a solid copper shield 0.0012 cm thick on one side, with a shielded round wire (Mil-W-16878). The average mass saving over round wire in the range AWG No. 20 to No. 32 is 78 percent. By using flat cables with shields on both sides for low frequencies, the average saving will be 74 percent.

Shields for higher frequencies generally need more distance between the conductor and the shield to keep capacitance within allowable limits. This additional dielectric material reduces the amount of weight saving.

TABLE V. MASS CHART FOR SHIELDED WIRE

Flat Wire (Mylar Insulation, 1 Side Shielded)						Round Wire (Stranded) Mil-W-16878		
Cond. Spacing 10^{-3}In	Cable Thickness 10^{-3}In	Cond. Dimen. 10^{-3}In	Cond. Area 10^{-6}In^2	Ohm m	Mass g/m	AWG	Area 10^{-6}In^2	Mass g/m (Average)
50	6.5	2 x 25	50	.547	.66	32	50	5.5
	7.5	3	75	.365	.82	30	79	5.8
	8.5	4	100	.270	.99	29	100	6.1
	9.5	5	125	.218	1.16	28	126	6.3
75	6.5	2 x 25	50	.547	.87	32	50	5.5
	7.5	3	75	.365	1.05	30	79	5.8
	8.5	4	100	.270	1.24	29	100	6.1
	9.5	5	125	.218	1.43	28	126	6.3
75	6.5	2 x 40	80	.337	1.01	30	79	5.8
	7.5	3	120	.227	1.27	28	126	6.3
	8.5	4	160	.151	1.53	27	158	7.4
	9.5	5	200	.136	1.79	26	200	8.2
100	6.5	2 x 65	130	.209	1.46	28	126	6.3
	7.5	3	195	.139	1.87	26	200	8.2
	8.5	4	260	.104	2.27	25	252	8.5
	9.5	5	325	.084	2.68	24	320	8.9
	10.5	6	390	.070	3.08	23	400	11.8
150	6.5	2 x 115	230	.118	2.36	25	252	8.5
	7.5	3	345	.079	3.05	24	320	8.9
	8.5	4	460	.059	3.75	22	510	15.8
	9.5	5	575	.047	4.43	21	640	17.6
	10.5	6	690	.039	5.13	21	640	17.6
	11.5	7	805	.034	5.82	20	800	20.0
	12.5	8	920	.030	6.51	19	1000	23.9

COST COMPARISON

The estimated cost figures in the following chart are based on a 5-foot, 25 conductor, round wire cable with soldered joints and a 5-foot flat conductor cable with molded-on plugs. The indicated 80 percent saving may be reduced to 70 percent by using crimp joints for the round wire cable.

Table VI shows only the comparison of purchase cost. It should be remembered that increased payload capability also affects the cost picture.

TABLE VI. COST CHART (Tentative Estimates)

	Round Wire Cable		Flat Cond. Cable	
	Material	Labor	Material	Labor
2 sockets	$16.00	$36.00	$16.00	
2 plugs	$16.00	$36.00	$ 1.00	$1.00
5-ft cable	$ 1.00	$12.00	$ 5.00	$1.00
Cost	$33.00	$84.00	$22.00	$2.00
Total Cost	$117.00		$24.00	
Saving*			80%	

*The cost saving for longer cables made of hook-up wire will be less.

NOTE:

The 80 percent cost saving shown in Table VI might be misleading if it is not carefully applied. The application of flat conductor cables for the entire spacecraft still requires much development work of shielded cables, power cable, connector hardware, and experience in installation. It also requires design consideration, careful planning, and much education of designers and craftsmen to materialize the full benefit of the flat conductor cable system.

RELIABILITY

The manufacture of flat conductor cables has improved considerably during the past few years. Most of the cables are manufactured by applying either heat-sealing or heat-curing plastics, making the plastic around all the conductors one seamless envelope. Tests prove that, at temperature extremes, round wire bundles are inferior to flat conductor cables in their ability to withstand flexing.

Another reliability factor is the number of junctions, which has been reduced in flat cable systems. Separable connections will usually have two junctions for the flat cables and at least three junctions for the round wire cables.

Because the flat conductors are cemented between plastic films forming a band or tape, no single conductor can be strained. This could easily happen in a hand-laced round wire cable, where the individual wires are not bonded to form a unit.

Reliability is improved and can be much better controlled by using machine tools or automation rather than handwork. Round wire harnesses and the addition of connectors are mostly handwork; miswirings are easily possible and costly to correct. All critical operations in making flat conductor cable harnesses are performed by machine tools, and handwork is almost eliminated; therefore, reliability of flat wiring is much greater than that of round wiring.

Special emphasis is made on the reliability of the connector design so that the final product is controlled not so much by the quality of manufacture, as by the principle of design and the establishment of proper and convenient manufacturing tolerances. This does not mean that control in manufacturing should be relaxed; it simply means that the principle of reliability is designed and not manufactured into the connector and is therefore inherent.

SPECIFICATIONS AND STANDARDS

The need for standardization was recognized early and committees from the Institute of Printed Circuits, Aerospace Industries Association, and Marshall Space Flight Center worked out specifications and standards for unshielded cables (Refs. 1 and 2). The cables are classified by operating temperature. Standards have been established for dimensions. The center spacings are arranged in multiples of .025 inches and the conductor sizes are, to a certain degree, selected to compare with AWG numbers.

PRODUCTION CAPACITY AND CONNECTOR DEVELOPMENT STATUS

Cable manufacturing development started several years ago. Much progress has been made in selecting the proper materials and procedures. Most companies have designed and built their own laminating machines.

Some of the equipment is for manufacturing studies and small quantities; other equipment is designed for production speed. Most companies are carefully protecting their proprietary rights, while others very openly show their laminating machines and discuss their problems. The lead time for cable delivery is relatively short.

Connectors for flat conductor cables have been studied and considered by many of the connector companies. Some have designed a whole family of connectors; others have built only prototypes for testing.

The Astrionics Laboratory has developed an environmental connector and manufactured a test series of 100 connectors, 15 of which have been tested by the Quality Assurance Laboratory. All connectors fulfilled the requirements of MSFC-SPEC-219. The tested type has 25 contacts for a 2-inch wide cable with 0.075-inch center spacing. Connectors for 1-, 1 1/2-, and 2-inch wide cables are available, and tooling for 2 1/2- and 3-inch cables is almost completed.

CONCLUSIONS

The highlights of the savings shown in detail are:

Connector Mass	25 - 75%
Cable Mass	25 - 78%
Material Cost	60%
Labor Cost	90%

These data are valid for designs using only the printed cabling method. They will be less favorable if adapters for equipment available only with round wire connectors must be used or if wiring changes become mandatory after manufacturing of the complete wiring as during research and development phases of a vehicle.

Flat conductor cables, both unshielded and shielded, are commercially available in various configurations. Specifications and standards for spacecraft environment have been established in cooperation with the Institute of Printed Circuits and the Aerospace Industries Association. Tooling for cable termination has been developed.

Further development is needed in the following areas:

(1) Development of different connector types such as solder pins, crimp-on pins, and wire-wrap pins. (A connector family of the plug-in type has been established.)

(2) Handtools for terminating cables inside of spacecraft. (Shop tools have been established.)

REFERENCES

1. MSFC-SPEC-219, Connectors for Flat Conductor, Flexible, Electrical Cable, October 3, 1962.

2. MSFC-SPEC-220, Cable, Electrical, Flat Conductor, Flexible, 300-Volt, June 20, 1963.

IV PUBLICATIONS, PRESENTATIONS, AND PATENTS...

A. PUBLICATIONS

SATURN I GROUND COMPUTER OPERATION, James L. Lewis, MTP-ASTR-N-63-14, September 19, 1963

A comprehensive reference source on the Saturn I ground computer interface hardware and operation with checkout systems of the Saturn I vehicle program is presented. Detailed explanations of the computer input-output assignments, characteristics, programing methods for access and control, and proposed system utilization are provided.

SATURN IB/V ASTRIONICS SYSTEM, S. M. Seltzer, MTP-ASTR-S-63-15, November 14, 1963

A brief, semi-technical, functional description of the astrionics systems for the Saturn IB and V launch vehicles is presented. The astrionics equipment consists of the vehicle guidance, control, and instrumentation systems. The adaptive guidance mode and the manner in which the mission is to take place are also briefly described.

The Saturn IB and V vehicles will use almost identical astrionics systems; minor differences may be caused by the reflection upon the astrionics system of two or three stages or by the assigned mission.

EFFECTS OF GAMMA RADIATION ON SELECTED POTTING COMPOUNDS AND INSULATING MATERIALS, Bobby W. Kennedy, MTP-ASTR-P-63-16, November 19, 1963

Recent radiation tests were performed at the Oak Ridge National Laboratory on the following types of polymeric materials: epoxy, silicone, and polyurethane resins; Teflon, polyester film, and Mylar flat-conductor cables. After exposure of up to 10 roentgens, mechanical properties were analyzed by the Pilot Manufacturing Branch. Property change data were then correlated with molecular effects caused by absorption of gamma rays. The effects of radiation on the physical properties of common polymeric materials that were irradiated at Oak Ridge are discussed.

At the maximum doses to which the subject material was exposed, Mylar and polyester films performed well. Teflon performed very poorly under the same conditions. Among the potting compounds, the epoxy resins showed the effects induced by both chain scission and crosslinking. The silicones changed only slightly in physical properties. Test data indicated a gradual degradation in the properties of polyurethane resins.

This report represents the initial portion of a continuing effort in the evaluation of polymeric materials for use in radiation environments. Present plans call for more extensive testing in combined space environments in the near future.

SPACE VEHICLE SA-3, TELEMETRY SYSTEM, W. B. Threlkeld, Jr., and E. H. Reeves, Jr., MTP-ASTR-I-63-17, November 26, 1963

The performance evaluation of the complete telemetry system (ten links) used for flight testing Saturn vehicle SA-3 is presented. The eight operable telemetry links, the two experimental telemetry links, and the telemetry (TM) auxiliary equipment assembly have been technically analyzed on an individual basis.

Statistical analyses were performed on much of the telemetry data and the results are presented.

The overall performance of the telemetry system used for flight testing SA-3 was as anticipated. It is inferred that the derived test data will facilitate the development of more efficient telemetry systems.

"Page missing from available version"

PAGES 124 – 125

CAM PRODUCING APPARATUS, Josef Boehm, Patent No. 3,106,868, October 15, 1963

This invention relates to an apparatus in which the cam cutting tool is positioned according to the master cam by means of a photocell controlled drive. Thus, the master cam is relieved from high contact pressures.

TAPE CABLES AND CORRESPONDING CONNECTOR, A. L. Herrmann, Patent No. 3,114,587, December 17, 1963

The ends of tape cable are mounted on pressure pads with raised areas. These areas lift the contacting parts of the cable conductors above the plane of the flexible cable. Details for removal of insulation and construction of a corresponding connector are given.